

Avoimet julkaisuarkistot ja niiden tekniikka

Mika Stenberg

Helsinki 9.4.2008

Pro gradu -tutkielma

HELSINGIN YLIOPISTO

Tietojenkäsittelytieteen laitos

Tiedekunta/Osasto — Fakultet/Sektion — Faculty		Laitos — Institution — Department	
Matemaattis-luonnontieteellinen		Tietojenkäsittelytieteen laitos	
Tekijä — Författare — Author Mika Stenberg			
Työn nimi — Arbetets titel — Title Avoimet julkaisuarkistot ja niiden tekniikka			
Oppiaine — Läroämne — Subject Tietojenkäsittelytiede			
Työn laji — Arbetets art — Level Pro gradu -tutkielma	Aika — Datum — Month and year 9.4.2008	Sivumäärä — Sidoantal — Number of pages 62 sivua + 3 liitesivua	
Tiivistelmä — Referat — Abstract <p>Open Access -liike pyrkii vapauttamaan tieteellisen tiedon kaupallisuuden rajoitteista edesauttamalla artikkeleiden rinnakkaisversioiden avointa ja esteetöntä verkotallennusta. Sen mahdollistamiseksi verkkoon perustetaan julkaisuarkistoja, joiden toiminta-ajatuksena on säilöä taustayhteisönsä tieteellinen tuotanto avoimesti ja keskitetysti yhteen paikkaan. Avoimen lähdekoodin arkistosovellukset jakavat sisältönsä OAI-protokollan avulla ja muodostavat näin globaalin virtuaalisen tietoverkon. Suurten tietomäärien käsittelyssä on huomioitava erityisesti kuvailutiedon rooli tehokkaiden hakujen toteuttamisessa sekä tiedon yksilöiminen verkossa erilaisten pysyvien tunnisteiden, kuten Handle:n tai URN:n avulla. Tieteellisen tiedon avoimella saatavuudella on merkittävä vaikutus myös oppimisen näkökulmasta. Julkaisuarkistot tarjoavat oppimateriaalin lisäksi uusia mahdollisuuksia julkaisukanavan ja oppimisympäristön integroimiseen.</p> <p>Työssä esitellään avoimen saatavuuden keskeisiä teemoja sekä sen käytännön toteutusta varten kehitettyjä teknisiä ratkaisuja. Näiden pohjalta toteutetaan Meilahden kampuksen avoin julkaisuarkisto. Työssä pohditaan myös julkaisuarkistojen soveltuvuutta oppimisprosessin tukemiseen tutkivan- ja sulautuvan oppimisen viitekehyksessä.</p> <p>ACM Computing Classification System (CCS): H.3 [INFORMATION STORAGE AND RETRIEVAL] H.3.7 [Digital Libraries], H.3.3 [Information Search and Retrieval], H.3.5 [Online Information Services] K.3 [COMPUTERS AND EDUCATION] K.3.1 [Computer Uses in Education]</p>			
Avainsanat — Nyckelord — Keywords Julkaisuarkistot, digitaaliset kirjastot, Open Access, DSpace, OAI, sulautettu oppiminen			
Säilytyspaikka — Förvaringsställe — Where deposited Kumpulan tiedekirjasto, sarjanumero C-			
Muita tietoja — Övriga uppgifter — Additional information			

Sisältö

1 Johdanto	1
2 Open Access – Avoin saatavuus	2
2.1 Määritelmä	2
2.2 Avoin Saatavuus käytännössä	3
2.3 Avoimuuden hyödyt	6
2.4 Haasteita avoimuuden etenemiselle	8
2.5 Open Access Suomessa	9
3 Avoimen julkaisemisen tekniikka	10
3.1 Yleistä	10
3.2 Yhteistoiminnallisuus ja roolit	11
3.3 Metatieto	12
3.4 OAI-protokolla	16
3.5 Pysyvät tunnisteet verkossa	18
3.6 Työnkulun hallinta	20
4 Julkaisuarkistot opetuksen ja oppimisen tukena	21
4.1 Oppimisen malleja	22
4.2 Julkaisuarkisto osana oppimisprosessia	24
5 Meilahden kampuksen avoin julkaisuarkisto	28
5.1 Käyttötarkoitus ja motivaatio	29
5.2 Sovelluksen valinta	30
5.3 Järjestelmän tekninen kuvaus	31
5.4 Järjestelmän tietosisältö	37

	iii
5.5 Peruskäyttö	41
5.6 Tiedon syöttöprosessi	44
5.7 Käyttötapaukset	48
5.8 Jatkokehitys	51
6 Pohdintaa	52
Lähteet	54
Liitteet	
1 CDSWaren ja DSpacen vertailu	
2 Julkaisuarkiston DSpace-käyttöliittymä	
3 Julkaisuarkiston Manakin-käyttöliittymä	

1 Johdanto

Ajankohtaisen tutkimustiedon saatavuus on tieteen tekemisen kulmakiviä. Sillä on merkittävä rooli myös opetuksen ja oppimisen tukena. Tieteelliset julkaisut ovat pitkään toimineet tiedon merkittävimpinä välittäjinä. Viimeisen kahden vuosikymmenen aikana tekniikan kehittyminen ja markkinavoimat ovat kuitenkin saattaneet perinteiset tiedonvälityksen mekanismit uusien haasteiden eteen.

Taloudellista voittoa tavoittelevien kustantajien mukaantulo on pikkuhiljaa johtanut tieteellisen tiedon kapitalisoimiseen; perinteinen tieteen tekemisen eetos, jossa tieto mielletään yhteiseksi hyväksi, on hämärtynt [Nii01]. Kustannustoiminnan kaupallistumisen [McC99, Wil06a, s. 17-18] ja lehtien tilaushintojen räjähdysmäisen kasvun myötä [Sub07, Wil06a, s. 15, 17, 26] tiedon saatavuus on heikentynyt, vaikka tekniikka sen jakamiseen kaikille halukkaille olisi mahdollista. Hinnoittelupolitiikasta johtuen tietoon pääsy on rajattu varakkaille, eivätkä hyvät ideat pääse kiertämään vapaasti. Nykyinen tieteellisen julkaisemisen malli, jossa tutkijoiden ilmaiseksi luovuttama tieto myydään kalliilla takaisin sen tuottamisen rahoittaneelle akateemiselle yhteisölle, on monien mielestä liian kallis ylläpitää [Wil06a, s. 18] ja kipeästi uusien julkaisumallien tarpeessa [ARL00, Wil06a, s. 25].

Näitä epäkohtia korjaamaan syntyi liike nimeltä Open Access (OA, suom. avoin saatavuus, avoin verkkojulkaiseminen). Nimensä mukaisesti sen kantavana ajatuksena on saattaa uusin tieteellinen tieto maksutta kaikkien saataville. Tähän pyritään perustamalla Internetiin avoimia julkaisuarkistoja ja tieteellisiä verkkolehtiä, joihin tutkimusartikkeleita voidaan tallentaa digitaalisessa muodossa [BOA02].

Tutkielmassani esittelen avoimen saatavuuden keskeisiä teemoja sekä sen käytännön toteutusta varten kehitettyjä teknisiä ratkaisuja. Näiden pohjalta hahmottelen avoimen lähdekoodin pohjalta rakennetun julkaisuarkistojärjestelmän, jossa toteutuu avoimen saatavuuden vaatimukset yhteiskäyttöisyydestä ja tiedon tavoitettavuudesta. Pohdin myös julkaisuarkistojen käyttömahdollisuuksia oppimisprosessin tukemiseen sulautuvan oppimisen viitekehyksessä.

2 Open Access – Avoin saatavuus

Tutkijat ovat kautta aikain pyrkineet maksimoimaan työnsä vaikutuksen ja laadun jakamalla artikkeliansa julkaisemattomia versioita kollegoilleen ilmaiseksi [Har06]. 1990-luvun alussa joukko tutkijoita alkoi samassa hengessä julkaista tutkimustuloksiaan Internetissä kaikkien saatavilla. Ennen pitkää näistä sivustoista muodostui tiiviisti seurattuja oman alansa ajankohtaisimman tiedon julkaisufoorumeita. Erityisen tunnetuiksi tulivat luonnontieteilijöiden ArXiv- ja Citeseer-palvelut, joihin kertyi nopeasti satojatuhansia dokumentteja. Huomattiin, että mahdollistamalla artikkeleiden avoin saatavuus – *Open Access* – ne leviävät ja tulevat nopeammin hyödynnetyksi kuin perinteisissä julkaisuissa, joiden rajoitteina ovat edelleen pitkät julkaisuajat sekä rajattu lukijakunta [Har06].

2.1 Määritelmä

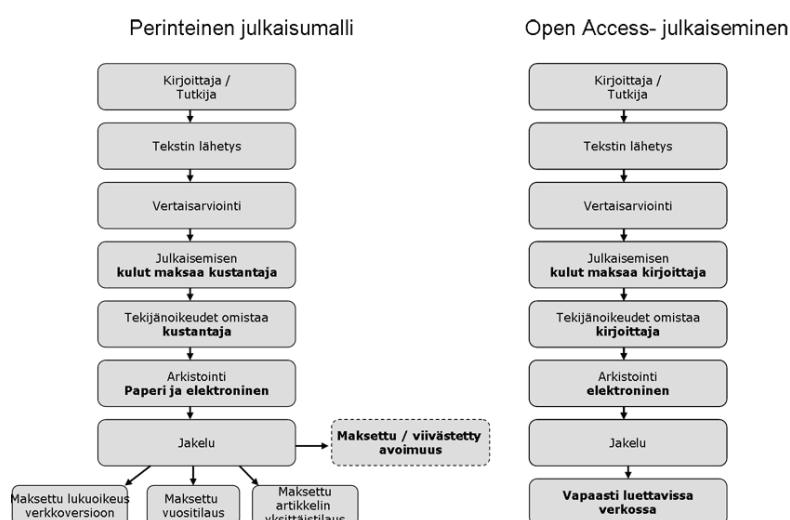
Open Access -käsitteen määritelmä kehittyä jatkuvasti käyttökontekstinsä mukaan. Ilmiön keskeinen sisältö on kuitenkin koottu joukkoon julkilausumia, joista kukin täydentää edeltäjiään. Kolmesta keskeisestä määrittelydokumentista käytetään syntypaikkojensa mukaisesti myös nimeä BBB -määritelmä¹ [Sub07, Bar06].

Nykyään avoimella saatavuudella tarkoitetaan tieteellisten artikkelien luonnosten sekä niiden julkaistujen versioiden rinnakkaiskopioiden saattamista ilmaiseksi kaikkien saataville Internetiin [Ope05]. Se voidaan nähdä myös uutena julkaisumallina, jossa tutkimusrahoittaja vastaa tutkimuksessa syntyneiden tulosten julkaisemisen kustannuksista [Coc06, Pro07, Ope05]. Näin tiedon loppukäyttäjä ei enää joudu maksamaan tiedon saatavuudesta, vaan ainoastaan sen painetusta versiosta (kuva 1). Avoimuuden vaatimuksen tekee mielekkääksi tieto siitä, että tutkijat eivät tavallisesti muutenkaan saa julkaisuistaan erillistä korvausta, vaan luovuttavat työnsä veloituksetta kustantajalle. Jakamalla artikkeleja ilmaiseksi verkon kautta vapautetaan kaupallisuuden luomista rajoitteista ja varmistetaan julkaisun hyödyllisyys sekä tutkimuksen tekijöille että koko akateemiselle yhteisölle unohtamatta satun-

¹Budapest, Bethesda, Berlin

naista tiedon etsijää [BOA02, Wil06a]. Samalla uusi julkaisumalli voi olla kuromassa umpeen digitaalisen kuilun aiheuttamaa eriarvoisuutta tiedon ja innovaatioiden jakamisessa länsi- ja kehitysmaiden välillä [Saa07]. Toisin kuin joskus on ajateltu, tutkimustiedon avoimuus ei uhkaa tiedon luotettavuutta, vaan toimii ennemminkin sen laadunvarmistajana tiedeyhteisön aktiivisesti arvioidessa uuden tiedon oikeellisuutta [Ket07].

Vaikka liikkeen alkuperäinen mielenkiinnon kohde olivat nimenomaan kokotekstimuotoiset tutkimusartikkelit, voidaan termi nykyään liittää lähes kaikkeen digitaaliseen aineistoon, jota tiedeyhteisöissä tuotetaan². Tyypillisiä esimerkkejä näistä ovat opinnäytteet, seminaari- ja konferenssiesitykset sekä kuvat ja videot [Bar06, DC07]. Viime aikoina myös oppimateriaalia sekä tutkimusaineistojen raakadataa on alettu tuoda liikkeen piiriin yhä enenevässä määrin.



Kuva 1: Julkaisumallit rinnakkain.

2.2 Avoin Saatavuus käytännössä

Määritelmän mukaan avoin saatavuus toteutuu kahdella toisiaan täydentävällä tavalla; puhutaan kahdesta polusta vapaaseen tietoon [HBV⁺04, BOA02]. Kultainen

²ns. gray literature

polku toteutuu, kun vapaa pääsy artikkeleihin tarjotaan kaikille avoimien tieteilisten verkkolehtien kautta. Vihreää polkua noudattavat tutkijat puolestaan itsearkistivat töitään, eli tallentavat omat julkaisunsa elektronisessa muodossa kaikkien saataville Internetiin. Sijoituspaikkana voi toimia tutkijan oma kotisivu tai jokin avoin julkaisuarkisto. Julkaisuja voi sijoittaa useampaankin paikkaan, mikä takaa niiden paremman saatavuuden jatkossa [Bar06].

Julkaisuarkistot

Julkaisuarkistot toimivat tietovarastoina tieteellisten artikkelien digitaalisille versioille tai muulle aineistolle, joka katsotaan sinne sopivaksi [DC07, JAM06, s. 10]. Arkistot muodostavat digitaalisia kirjastoja, joita kutsutaan käyttökontekstinsa mukaan laitoskohtaisiksi (institutional repository) tai aihekohtaisiksi julkaisuarkistoiksi (subject-based archives). Arkistoissa materiaali järjestetään perinteisten kirjastojen tapaan hierarkkisiin kokoelmiin. Niiden nimeämiskäytännöt ja materiaalin ryhmittely kuitenkin vaihtelevat suuresti. Rakenne määrittäneekin yleensä julkaisuarkistoa hallinnoivan tahon taustasitoumusten kautta. Tavallisesti laitoskohtaiset julkaisuarkistot ryhmittelevät aineistonsa tiedekunta- ja laitosrajojen mukaisesti. Aihekohtaisten arkistojen ryhmittely sen sijaan noudattelee erityisaloja ja tarkemmin rajattuja teemoja. Sisältöpolitiikka siis profiloi julkaisuarkistoja voimakkaasti [JAM06, s. 6-8][BW05, s. 19].

Vaikka molemmat edellä mainitut mallit palvelevat tieteen avoimuuden toteutumista, jakavat ne voimakkaasti mielipiteitä puolesta ja vastaan. Erikoisalojen ympärille rakentuneiden arkistojen on ajateltu keräävän enemmän materiaalia ja ylittävän näin helpommin sen kriittisen massan, joka vaaditaan palvelun elinehdoksi. Ne ovat myös löyhemmin hallinnoituja ja lähteneet liikkeelle nimenomaan tutkijoiden omasta aloitteesta, ilman hallinnollista painetta, mikä viestii sisäsyntyisestä innostuksesta asiaan [Cro02, s. 21]. Laitosmallin kannattajien mukaan yksikkötasolla tapahtuva arkistointi mahdollistaa selkeämmän sisältöpolitiikan sekä paremmat mahdollisuudet tieteellisen tuotannon järjestelmälliseen tallentamiseen [SB05, s. 5-6][Ope05, s. 22]. Jonkin erikoisalan ympärille voidaan tämän jälkeen helposti luoda omia palveluitaan

kokoamalla aihealueittain valikoitua metatietoa keskitetysti yhteen. Yliopistojen ja tutkimuslaitosten perustaessa yhä kiihtyvällä tahdilla omia julkaisuarkistojaan vaikuttaisi siltä, että laitoskohtainen arkistomalli koetaan yleisesti toimivammaksi.

Maaailmanlaajuisesti julkaisuarkistoja on toiminnassa jo lähes tuhatkunta. Näistä valtaosa sijaitsee Yhdysvalloissa. Julkaisuarkistoista pidetään listaa mm. OpenDoar³ sekä ROAR⁴ -projektien kotisivuilla.

Avoimet tieteelliset verkkojulkaisut

Avoimella verkkojulkaisulla (Open Access journal) tarkoitetaan verkossa ilmestyvää, tieteelliset normit täyttävää ja vertaisarvioitua verkkolehteä, jonka sisältämiin kokotekstiartikkeleihin kaikilla on vapaa pääsy. Tilausmaksujen sijaan julkaisemisen ja laadunvalvonnan kulut pyritään kattamaan sponsorirahalla, vuosijäsenmaksuilla tai laskuttamalla artikkelin kirjoittajan taustayhteisöä julkaisuprosessin yhteydessä (kuva 1). Toisin kuin monissa kaupallisissa lehdissä, avoimissa verkkojulkaisuissa artikkelien tekijänoikeudet jäävät kirjoittajille [Ope05, Sub07].

Avoimuuden vaatimuksen tullessa yhä voimakkaammaksi myös monet kaupalliset kustantajat ovat alkaneet tarjota mahdollisuutta julkaista avoimia artikkeleita lisämaksua vastaan. Tästä käytetään nimitystä maksettu avoimuus; Paid Open Access⁵. Loppukäyttäjälle tällainen palvelu näyttäytyy hybridijulkaisuna, jossa maksullisten artikkelien joukossa on myös kaikille avoimia dokumentteja [Bar06]. Jotkut julkaisijat ovat ottaneet käyttöönsä myös ns. viivästetyn Open Access -mallin. Siinä maksavien asiakkaiden etuja varjellaan vapauttamalla artikkeli automaattisesti kaikkien luettavaksi 6-12 kk:n viiveellä [Wil06a, s. 118, 212].

Lundin yliopiston kirjasto ylläpitää kattavaa hakemistoa avoimista tieteellisistä verkkojulkaisuista. Directory of Open Access Journals⁶ -sivustolla on listattu noin 2000 lehteä, joiden sisältö on vapaasti kaikkien luettavissa. Toinen merkittävä hakemisto

³<http://www.opendoar.org>

⁴<http://roar.eprints.org>

⁵Esimerkiksi Springer Open Choice tarjoaa tällaista palvelua

⁶<http://www.doaj.org>

on Brasiliassa ylläpidetty SciELO⁷, joka sisältää noin 300 espanjan- ja portugalin-kielistä tieteellistä julkaisua.

2.3 Avoimuuden hyödyt

Tutkija ja kustantaja

Tutkijoiden kannalta kiinnostavin avoimuuden tuoma hyöty lienee sitaattimäärien lisääntyminen. Sen seurauksena myös julkaisujen arvostuksesta kertova vaikuttavuuskerroin nousee [SB05].

Vertailtaessa artikkeleihin kohdistuneita viittausmääriä huomattiin, että julkaisut joista oli saatavilla myös kaikille avoin versio saivat keskimäärin kaksinkertaisen määrän viittauksia muihin nähden. Verkosta löytyviin vanhempiinkin artikkeleihin viitataan ahkerasti, mikä pidentää niiden elinkaarta [KB06]. Tieteenalakohtaisista eroavaisuuksista on havaittavissa, mutta ainakin tietojenkäsittelytieteessä [Law01], hiukkasfysiikassa [Bro04], avaruusfysiikassa [KB06] sekä monilla muilla tieteenaloilla [HSG05] vaikutukset ovat olleet positiivisia. Tähtitieteessä avoimeen verkkojulkaisuun siirtymisestä ei kuitenkaan havaittu koituvan etua. Tämän voidaan katsoa olevan erikoistapaus, sillä alan tutkijoiden joukko on melko pieni ja lähes kaikilla on pääsy muutamiin sen keskeisiin julkaisuihin [KEA⁺05].

Jos sitaattien määrää verrataan eri lehtien kesken, jää avoimien lehtien etu hyvin pieneksi [Pri04]. Ottaen kuitenkin huomioon, että lehden kehittyminen tunnustetuksi ja arvostetuksi toimijaksi alalla vaatii aikaa, on yllättävää, että uudehkot avoimet lehdet pärjäävät hyvin vertailussa tilausperustaisia lehtiä vastaan [KB06].

Avoimesti saatavilla olevien artikkelien vaikuttavuuden lisääntyminen perustuu niiden saatavuuden parantumiseen ja julkaisun nopeutumiseen. Avoimesti verkkoon tallennettu artikkeli on kaikkien saatavilla välittömästi, ilman kalliita tilausmaksuja. Tutkimuksen kierto lyhenee eikä uuden tiedon julkistamista tarvitse odottaa painetun lehden ilmestymiseen saakka [SB05, s. 5] Jonkin verran vaikutusta näyttäisi olevan myös sillä, että kirjoittajat julkaisevat mieluiten juuri parhaina pitamiään

⁷<http://www.scielo.org>

artikkeleita avoimesti verkossa (ns. quality bias) [KB06].

Hallinto

Korkeakouluille ja tutkimuslaitoksille julkaisujen tuominen verkkoon tarjoaa mahdollisuuden edistää tehdyn tutkimuksen näkyvyyttä ja vaikuttavuutta sekä kartoittaa tieteellisen julkaisemisen kokonaistilannetta yksikkötasolla. Julkaisuarkistoa voidaan käyttää tehdyn tutkimuksen markkinointitarkoituksiin sekä erityisosaamisen esiintuomiseen. Arkistot tarjoavat karkean vertailumittarin muualla tehtyyn tutkimukseen sekä laadullisesti että määrällisesti ja mahdollistavat tieteellisen tuotannon kokonaisvaltaisen arkistoinnin ja pitkäaikaissäilytyksen [Swa07, DC07, SB05, s. 5]. Hallinnon tasolla ollaan puolestaan erityisen kiinnostuneita yksiköiden tuottavuudesta ja sen seurannan työkaluista, joiden pohjalta kerättyä tietoa voidaan käyttää myöhemmin laadun arvioinnissa ja rahanjaossa. Arkistoihin keskitetty julkaisujärjestelmä mahdollistaa uusia tapoja julkaisujen tilastointiin ja niiden käytön seurantaan [MB07]. Niiden kautta voidaan keskitetysti hallita myös tutkimushankkeita ja niihin kuuluvia resursseja laajemminkin. Tällaista järjestelmien integrointia ollaankin suunnittelemassa mm. Helsingin yliopiston julkaisuviitetietokannan (Julki) ja tutkimuksenhallintajärjestelmän (Tuhat) sekä kampusarkistojen yhteyteen.

Oppiminen

Laadukkaan opetusmateriaalin lisäksi todellinen oppiminen vaatii myös opetusmateriaalin ulkopuoliseen, ajankohtaiseen tutkimustietoon perehtymistä [Arm00]. Avoimesti saatavilla olevat tieteelliset julkaisut muodostavat erinomaiset resurssit tätä varten ja tuovat tieteellisen kirjaston helposti kaikkien ulottuville [MM95]. Sitä hyödyntävät tutkijoiden lisäksi myös opiskelijat ja opettajat sekä yhteiskunnan päättäjät [Wil06b, Wil06a, s. 88, 110-111]. Tieteellisen tiedon avoin saatavuus verkossa muodostaa perustan elinikäiselle oppimiselle, jossa kellä tahansa tietoa kaipaavalla on resurssit sen hankkimiseksi.

2.4 Haasteita avoimuuden etenemiselle

Avoimen saatavuuden toteutumisen haasteista vain murto-osan arvioidaan liittyvän tekniikkaan, valtaosan ollessa hallinnollisia, rakenteellisia ja käytäntöihin liittyviä [Jak07, Ope05]. Myös tieteenalakohtaisilla julkaisukäytännöillä näyttäisi olevan keskeinen merkitys siihen, koetaanko avoin julkaiseminen toimivaksi vaihtoehdoksi. Luonnontieteissä pitkät perinteet, artikkelien julkaisemisen kiivas tahti sekä tutkimustiedon luonne, unohtamatta kaupallisuutta, ovat osoittautuneet sille otolliseksi maaperäksi. Sen sijaan humanistisissa tieteissä alan tutkijoiden mielenkiinto on vasta hiljalleen heräämässä, mikä johtunee kirjoihin ja monografioihin keskittyneestä kustannustoiminnasta sekä tiedon hitaammasta kierrosta [Cro02, s. 21][DC07, Lom06, s. 5-6].

Materiaalin hankinta

Merkittävimmät haasteet tiedon avoimuudelle asettaa itsearkistointikäytännön jalkauttaminen tutkijayhteisön keskuuteen ja sitä kautta materiaalin hankinta olemassa oleviin järjestelmiin [SB05]. Swanin & Brownin tekemän selvityksen mukaan vain noin kolmannes tutkijoista tuntee Open Access -liikkeen ja heilläkin esiintyy usein virheellisiä käsityksiä sen sisällöstä. Itsearkistointia kertoo tällä hetkellä harrastavansa noin 20 % tutkijoista, joskin 49 % aikoo seuraavan 3 vuoden aikana tehdä niin. Valtaosa vastaajista olisi kuitenkin valmis siihen työnantajan niin vaatiessa. Epäselviksi koetaan itsearkistointiin liittyvät tekijänoikeudelliset asiat, prosessin viemä aika sekä siihen liittyvät käytännön vaikeudet [Swa06, SB05, s. 56-57]. Myös plagiointia pelätään, vaikka keskitetyt dokumenttiarkistot itse asiassa mahdollistavat vilpin havaitsemisen entistä tehokkaammin [SGWG06]. Näyttäisikin siltä, että laajasta tiedotuksesta ja akateemisen yhteisön herättämästä keskustelusta huolimatta avainasemassa olevia tutkijoita ei ole vielä kunnolla tavoitettu [Ope05, SB05]. Laitosten tulisi panostaa uusien sisällönhankintastrategioiden kehittämiseksi ja toimeenpanemiseksi [JAM06, s. 111)].

Kokemukset maailmalta osoittavat, että nopein ja tehokkain tapa aktivoida tutkijat ja kerryttää materiaalia arkistoihin on ottaa käyttöön työnantajan tai tutkimuksen

rahoittajan taholla tehdyt auktoriteettipäätökset ja liittää näin töiden arkistointi osaksi laitosten normaalia tutkimus- ja julkaisukäytäntöä [Har06, Sal06, Jak07, JAM06, s. 135-136]. Helsingin yliopistossa ainakin eläinlääketieteellinen tiedekunta on lisännyt toimintakäsikirjaansa maininnan lisensiaatin tutkielmien säilömisestä elektroniseen arkistoon⁸.

Julkaisemisen rahoitus

Julkaisijoille tiedon avoimuus voi merkitä tilausmaksujen menetystä ja taloudellisia vaikeuksia. Näitä voitaneen kuitenkin kompensoida erilaisilla avoimuuden perusteella myönnettyillä julkaisutuilla [Ket07, Ope05, s. 17]. Avoimuus voi myös houkutella lisää julkaistavia käsikirjoituksia sekä kiinnostuneita lukijoita. Esimerkiksi fysiikan alan verkkojulkaisujen sisältö on ollut kaikkien saatavilla jo vuodesta 1991, eikä vaikutuksia paperiversioiden tilausmääriin ole havaittu. Rinnakkaiselo kahden julkaisutahon kanssa nähdään positiivisena asiana, jossa voidaan tehdä yhteistyötä [SB05, s. 4].

2.5 Open Access Suomessa

Suomessa avoin verkkojulkaiseminen on saanut jalansijaa vasta muutamina viime vuosina, kun yliopistot ja korkeakoulut yhdessä kirjastojen kanssa ovat alkaneet perustaa julkaisuarkistoja tutkimusartikkelien ja opinnäytteiden säilömistä varten. Merkittävä askel asian edistämiseksi on ollut opetusministeriön julkaisema Avoin Tieteellisen Julkaisutoiminnan työryhmän suositus vuoden 2005 alkupuolella. Siinä otettiin kantaa avoimen verkkojulkaisun puolesta ja suositeltiin konkreettisia toimia sen tukemiseksi [Ope05]. Kesällä 2006 Suomen yliopistojen rehtorien neuvosto allekirjoitti Berliinin julkilausuman tieteellisen tiedon avoimesta saatavuudesta⁹. Kansainvälisellä tasolla sekä Euroopan Unionin raportti tieteellisestä julkaisemisesta että YK:n tietoyhteiskuntakokous suosittavat tieteellisen tiedon avoimuutta omista

⁸Osa B2, s. 14(23) http://www.vetmed.helsinki.fi/laatu/toimintakasikirja_20080131.pdf

⁹<http://www.openaccess.fi/info/may06.html>

raporteissaan [EC06, YK03].

Käytännön tasolla asiaa tekee tunnetuksi FinnOA -työryhmä. Toiminnallaan se pyrkii edistämään julkaisuarkistojen perustamista ja tieteellisten lehtien siirtymistä avoimeen julkaisuun. Toiminta kulminoituu OA-JES -projektiin, jonka puitteissa koordinoidaan kansallista julkaisuarkistotyötä, järjestetään tiedotustilaisuuksia sekä vaihdetaan kokemuksia eri laitosten ja yksiköiden välillä¹⁰. Tästä näkyvimpänä osoituksena ovat Helsingin yliopiston neljä kampusarkistoa sekä Kansalliskirjaston e-thesis- ja Doria-verkkopalvelut. Suomi on myös aktiivisesti mukana pohjoismaisessa yhteistyössä.

Syyskuussa 2007 kotimaisia julkaisuarkistoja oli kymmenkunta, joista lähes puolet toimivat Helsingin yliopiston kampusarkistoina. Monet muut tahot, kuten ammattikorkeakoulujen konsortio sekä Svenska handelshögskolan (Hanken) ovat parhaillaan perustamassa omia arkistojaan. Avoimia tieteellisiä verkkolehtiä ilmestyy Suomessa tällä hetkellä kymmenisen kappaletta, osa niistä on saatavissa myös maksullisina painettuina versioina [Ope05, s. 19].

3 Avoimen julkaisemisen tekniikka

Avoimen verkkojulkaisemisen tekniset haasteet liittyvät ennen kaikkea tiedon elinkaaren hallintaan, monimuotoisen tiedon kuvailuun ja arkistointiin sekä tehokkaiden tiedonsiirto- ja hakutoimintojen tekemiseen yhden tai useamman tietojärjestelmän välillä. Teknisen rungon järjestelmille muodostavat Internet ja korkeakoulujen olemassa olevat tietotekniikkaympäristöt, jotka jo nyt yhdistävät tutkimuslaitoksia ympäri maailmaa.

3.1 Yleistä

Paikallisen julkaisuarkiston muodostaa palvelimelle asennettu ohjelmisto sekä siihen liittyvä tietokanta, joka pitää sisällään kuvailutietoja ja niihin liittyviä tiedostoja.

¹⁰<http://www.openaccess.fi>

Käyttäjilleen järjestelmä tarjoaa palveluita WWW-pohjaisen käyttöliittymän kautta. Järjestelmien alustana toimii yleensä joku useista verkossa saatavilla olevista avoimen lähdekoodin julkaisuarkisto-ohjelmistoista, kuten DSpace tai CDS Invenio. Perustoiminnallisuuden ohjelmissa muodostavat olemassa olevan tiedon selailu ja hakujen tekeminen järjestelmään. Lisäksi järjestelmän on kyettävä mukautumaan vaihteleviin tiedonsyöttöprosesseihin ja erilaisiin kuvailutietojoukkoihin käyttäjäyhteisönsä tarpeiden mukaisesti [JAM06, s. 85]. Näiden ohella järjestelmät tarjoavat usein erilaisia lisäpalveluja kuten materiaalin konversiota, kommentointityökaluja ja uutuuksseurantaa sähköpostilistojen tai RSS-syötteiden välityksellä. Pidemmälle vietyjä järjestelmän taustayhteisön tuottamia kaupallisia palvelukonsepteja voivat olla esim. artikkelin paperiversion tilausmahdollisuus tai kokoelmatilan vuokraaminen asiakkaiden materiaalia varten [GA08].

Ominaisuuksiensa perusteella julkaisuarkistosovellukset voitaneenkin jakaa karkeasti kolmeen ryhmään: tutkimus- tai oppimateriaalia varten tehdyt järjestelmät voidaan erottaa sisällönhallintajärjestelmistä [DeR04]. Keskeisiä ominaisuuksia kaikille ohjelmistoille ovat kattava tuki erilaiselle metatiedolle, joustavat työnkulun määrittelyt sekä tiedon arkistointi ja siirtomahdollisuudet eri järjestelmien välillä.

3.2 Yhteistoiminnallisuus ja roolit

Avoimet julkaisuarkistot eroavat perinteisistä kirjastojärjestelmistä yhteistoiminnallisuudellaan. Jakamalla tietonsa avoimesti kaikkien käyttöön, paikallinen julkaisuarkisto voi toimia rakennuspalikkana laajassa kansainvälisesti hajautetussa tietojärjestelmässä. Tällä tavoin ne muodostavat jättimäisen virtuaalisen tietovaraston, joka mahdollistaa materiaalin globaalin julkaisemisen ilman kaupallisia kustantajia ja tarjoaa samanlaisia palveluita ja käytettävyyttä kuin suuret kaupalliset järjestelmät [JAM06, s. 11][Cro02, s. 7]. Yhteistoiminnallisuuden tekee erityisen haasteelliseksi käytössä olevien järjestelmien heterogeenisuus; osajärjestelmät eroavat niin tekniikaltaan kuin sisällöltäänkin toisistaan huomattavasti [JAM06, s.11]. Tästä huolimatta niiden välinen tiedonsiirto on mahdollista käyttämällä yhteisesti sovittuja rajapintoja, kuten OAI-protokollaa.

Julkaisuarkistoilla on kahdenlaisia rooleja: ne voivat toimia palvelun tarjoajina (service provider) tai tiedon tarjoajina (data provider). Roolit voivat olla myös osittain päällekkäisiä riippuen halutusta toiminnallisuudesta sekä järjestelmän teknisestä arkkitehtuurista (kuva 2). Tiedon tarjoajaksi kutsutaan sellaista julkaisuarkistoa, joka tarjoaa sisältönsä osittain tai kokonaan kenen tahansa luettavaksi. Tällaisia järjestelmiä ovat pääasiassa yliopistojen ja tutkimuslaitosten kampusarkistot. Palvelun tarjoaja sen sijaan toimii viitetietokantana, johon kerätään metatietoa toisista julkaisuarkistoista [JAM06, s. 68]. Näin ollen se tarjoaakin yhteisen hakunäkymän joukkoon julkaisuarkistoja. Hakutulokset sisältävät kuvailutietojen lisäksi linkin artikkelin alkuperäislähteenä toimivaan julkaisuarkistoon, jossa säilytetään usein myös siihen liittyvää kokotekstidokumenttia. Tällaisia palveluja verkossa ovat esim. OAIs-ter¹¹, ARC¹² ja CiteBase¹³.

Edellä mainitut roolit voivat esiintyä myös yhdessä; arkisto voi toimia tiedon tarjoajana samalla kun se haravoi toisista arkistoista viitetietoja omiin kokoelmiinsa ja toimii näin palvelun tarjoajana. Esimerkiksi kampusarkistoihin voidaan tällä tavoin kerätä relevanttia metadataa toisista tietokannoista tai julkaisuluetteloista. Muualta tuotua viitetietoa voidaan jälkeinpäin helposti täydentää ja käyttää raakadatana syötettäessä uusia tiedostoja järjestelmään.

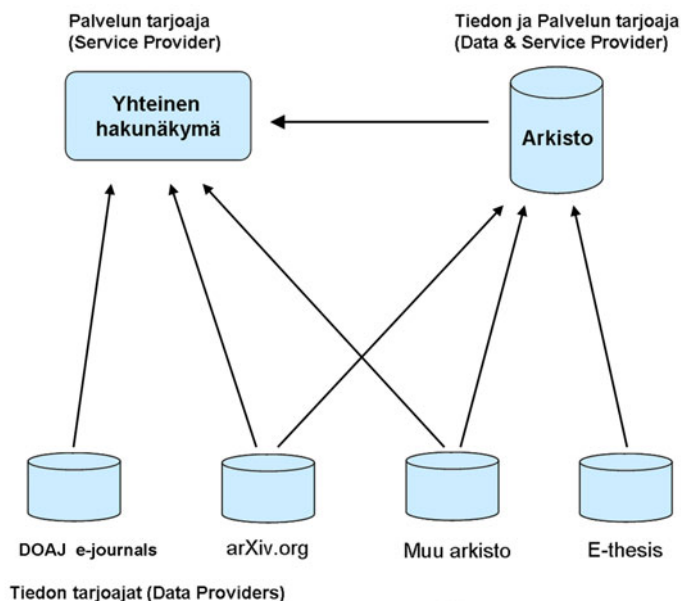
3.3 Metatieto

Metatieto on tietoa kuvaavaa ja määrittelevää tietoa. Suurten tietomäärien järjestämisessä ja tehokkaiden hakujen toteuttamisessa sillä on merkittävä rooli. Laadukkaan kuvailutiedon arvo onkin suunnaton; sen tuottaminen vie aikaa ja vaatii asiantuntemusta. Siitä on tullut kauppatavaraa, jota suuret kustannustalot myyvät kirjastoille.

¹¹<http://oaister.umd.umich.edu/o/oaister/>

¹²<http://arc.cs.odu.edu/>

¹³<http://www.citebase.org/>



Kuva 2: Julkaisuarkistoja erilaisissa rooleissa ja tiedon siirtyminen niiden välillä.

Yleistä

Metatieto koostuu joukosta ennalta määritellyjä elementtejä, joiden avulla materiaaliin liitetään kuvailutietoja. Perinteisten tekijä- ja julkaisutietojen lisäksi materiaaliin voidaan liittää tietoa esim. sen tallennusformaattista, käyttöoikeuksista, tekijänoikeussuojasta tai suunnitellusta kohdeyleisöstä. Yhteiskäytön tehostamiseksi tieto esitetään usein XML-muodossa. Kuvailutiedon rakenteisuus onkin avain haku- ja tarkkaan kohdentamiseen; se mahdollistaa haun kohdentamisen kenttätasolla. Se myös lisää tiedon uusiokäytön mahdollisuuksia ja luo puitteet datan hyödyntämiseen muissa järjestelmissä.

Kuvailutieto voidaan liittää osaksi kuvailtavaa dokumenttia (embedded) tai se voi sijaita erillisenä tiedostona (external) ja tarjota linkin kuvailemaansa materiaaliin. Kuvailutiedon sisällyttäminen binäärimuotoisiin tiedostoihin onkin pitkään ollut haasteellista [Hak02, s. 114, 142]. Tilanne on kuitenkin muuttumassa uusien XML-pohjaisten tiedostoformaattien myötä. Tämä luonee aivan uusia mahdollisuuksia metadatan sisällyttämisessä osaksi asiakirjoja.

Dublin Core

Julkaisuarkistoissa yleisesti käytetty Dublin Core -formaatti pyrkii yhdistämään yksinkertaisen kuvailujoukon ja ilmaisuvoimaisen kenttärakenteen yhdeksi paketiksi [Hak02]. Siinä ei ole varsinaisia kuvailusääntöjä, vaan jokainen määrittelee itse tavallaan, jolla soveltaa kielen tarjoamaa kenttävalikoimaa omassa käytössään (kuva 3). Kuvailuelementtejä on 15 kpl, joista kaikkiin on myös mahdollista liittää joukko erilaisia tarkentimia (qualifiers). Näiden avulla vakiokenttien käyttömahdollisuuksia voidaan laajentaa tuntuvasti¹⁴. Käyttöön voidaan ottaa myös erilaisia laajennuspaketteja (extension), kuten vielä kehitteillä oleva DC Educational Extension, joka esittelee oppimateriaalin kuvailuun tarvittavia lisäkenttiä. Tarkentimien käyttö jakaa kuitenkin mielipiteitä; toiset haluavat minimoida niiden käytön ja pitää kuvailujoukot yksinkertaisina, toisten mielestä niiden käyttö puolestaan rikastaa kuvailua [Hak02].

Dublin Core on otettu ristiriitaisesti vastaan. MARC:in¹⁵ kaltaisiin kieliin tottuneet kritisoivat sitä, kun taas monet muut pitivät sen käyttöönottoa tervetulleena yksinkertaistuksena verkkomateriaalin kuvailuun, jossa perinteiset menetelmät eivät materiaalin luonteen vuoksi ole enää käyttökelpoisia. Merkittävimmit haasteet DC:n käytössä liittyvät aihe- ja identifikaatiotunnus-kenttien käyttöön. Tämä ei johdu niinkään formaatista, vaan kenttien harkitsemattomasta ja epämääräisestä käytöstä [Hak02]. Aineiston identifiointia puolestaan vaikeuttaa verkkoaineiston erityisluonne sekä Internetin alati muuttuvat URL-osoitteet [Hak02, Kes07]. Huomiota on kiinnitetty myös Dublin Coren siirrettävyyteen; sitä on helppo konvertoida muille metakielille, mutta laajempien kuvailukielten palauttaminen siihen saattaa tuottaa vaikeuksia riittämättömien kenttämäärittelyjen vuoksi.

Vaikka Dublin Core asettaa rajoitteita bibliografia-tasoiseen tiedonkuvailuun, siitä on tullut merkittävä työkalu OAI-protokollaa käyttävien tietojärjestelmien välisessä kuvailutiedon vaihdossa. Se voidaan nähdä pienimpänä yhteisenä nimittäjänä erilaisten kuvailutieto-skeemojen joukossa, joka julkaisuarkistokäytössä tarjoaa käyttä-

¹⁴<http://www.dublincore.org>

¹⁵Machine Readable Cataloguing

jälle linkin takaisin materiaalin alkuperäislähteelle ja sen myötä myös kattavampaan kuvailutietoon [JAM06, s. 69].

```
<dublin_core schema="dc">
  <dcvalue element="title" qualifier="none">Vitamin C for preventing and treating the common cold </dcvalue>
  <dcvalue element="contributor" qualifier="author">Douglas, Robert M.</dcvalue>
  <dcvalue element="contributor" qualifier="author">Hemilä, Harri</dcvalue>
  <dcvalue element="identifier" qualifier="citation">PLOS medicine: 2 (2005) : 6, s. e168, [2 s.].</dcvalue>
  <dcvalue element="identifier" qualifier="issn">1549-1676</dcvalue>
  <dcvalue element="identifier" qualifier="uri">http://hdl.handle.net/2455/127938</dcvalue>
  <dcvalue element="format" qualifier="medium">elektroninen aineisto</dcvalue>
  <dcvalue element="language" qualifier="iso">eng</dcvalue>
  <dcvalue element="type" qualifier="none">artikkeli kausijulkaisussa</dcvalue>
  <dcvalue element="identifier" qualifier="JULKI">B3</dcvalue>
  <dcvalue element="identifier" qualifier="LAITOSKOODI">307</dcvalue>
  <dcvalue element="identifier" qualifier="pubmedid">15971944</dcvalue>
</dublin_core>
```

Kuva 3: Artikkelin metatiedot Dublin Corella kuvailtuna.

Moniulotteinen metadata

Kuvailtavien objektien muuttuessa yhä moniulotteisemmiksi on kehitetty uusia kuvailutietoskeemoja vastaamaan niiden asettamia vaatimuksia. Tällaisia ovat mm. METS, MODS ja SCORM. Ne mahdollistavat usean erilaisen kuvailuskeeman pake-toimisen yhdeksi kokonaisuudeksi. Näin dokumentin metatieto jakaantuu osiin käyt-tötarkoituksensa mukaisesti. Oman kokonaisuutensa voivat muodostaa esim. hallin-nollinen, tekninen, rakenteellinen ja tekijänoikeudellinen metatieto. Varsinainen si-sältöä kuvaileva metatieto muodostaa siis vain yhden osan valtavassa tietomassassa, joka elektronisiin dokumentteihin on liitettävissä [JAM06, s. 94-95].

Kuvailun ja metatiedon hyvistä käytännöistä on julkaistu suosituksia myös kotimai-seen käyttöön. Tällaisia ovat esim. Kansalliskirjaston ehdotus opinnäytteiden ku-vailua varten¹⁶ sekä Julkisen Hallinnon Tietohallinnon neuvottelukunnan suositus asiakirjojen kuvailuun¹⁷. Suomalaisen oppimateriaalin kuvailua varten ollaan myös määrittelemässä omaa kuvailuskeemaa, joka kulkee nimellä FinnEduMeta¹⁸.

¹⁶http://ethesis.helsinki.fi/metadata/metadata1_0.pdf

¹⁷<http://www.jhs-suositukset.fi/suomi/jhs143>

¹⁸http://www.virtuaaliyliopisto.fi/data/vyscripts/kokousymparisto_main/getdocument.php?doc=7007&PHPSESSID=0818a3e23dd26ef5c229cbcb34279669

3.4 OAI-protokolla

Julkaisuarkistojen välinen metatiedon siirto, nk. ”haravointi” (harvesting), tapahtuu erityisesti sitä varten kehitetyn OAI-protokollan (OAI-MHP, Open Access Initiative Metadata Harvesting Protocol) avulla. Se on yksinkertainen ja avoin tekninen standardi, ja määrittelee joukon toimintoja, joiden avulla järjestelmästä voidaan tuoda kuvailutietoja ulos XML-muodossa. Standardi määrittää pienimmän yhteisen nimittäjän, joiden varassa hyvinkin erilaiset julkaisuarkistot voivat kommunikoida keskenään ja mahdollistaa näin niiden kattavan yhteiskäytön [SW04, VdSNL04]. Alun perin julkaisuarkistojen tarpeisiin kehitetyn tekniikan on myöhemmin huomattu soveltuvan erinomaisen hyvin yleiskäyttöiseksi tiedonsiirtomekanismiksi erilaisten tietojärjestelmien välillä. Sen etuina ovat yksinkertaisuus ja tehokkuus sekä riippumattomuus sisällön tyypistä [Lyn01].

Protokolla käyttää yksinkertaista HTTP-pyynnön kysymys-vastaus -paria tiedon hakijan (haravoijan) ja tietovaraston väliseen kommunikointiin. Kohdejärjestelmän OAI-rajapintaan lähetetään tiedonhakupyynnöjä, jotka vastaanottaja tulkitsee ja suoritettuaan tarvittavat haut omassa tietojärjestelmässään palauttaa vastauksen. Vastauksen muodostaa XML-muotoinen esitys hakutuloksista, joka koodataan käyttäen pyydettyä kuvailukieltä [SW04, JAM06, s. 70-71]. Julkaisuarkistokäytössä hakutulokset esitetään tavallisesti Dublin Core -muodossa, mutta protokollan määrittely mahdollistaa myös muun tyyppisen metatiedon siirron. Tämä johtanee tulevaisuudessa monipuolisempien yhteisökohtaisten kuvailujoukkojen käyttöönottoon [Lyn01]. Vastauksen saatuaan kyselyn lähettäjä voi hyödyntää saamaansa hakutulosta parhaaksi katsomallaan tavalla; se voidaan esim. liittää osaksi paikallista julkaisuarkistoa tai siihen voidaan kohdistaa hakuja omana kokonaisuutenaan jossain toisessa sovelluksessa.

OAI tukee kaikkiaan kuutta erilaista toimintoa, joita kutsutaan verbeiksi (verb). Näiden avulla kohdejärjestelmästä voidaan kysyä erilaisia tietoja. Verbien toimintaa puolestaan tarkennetaan parametrien avulla, jotka kohdistavat toiminnot haluttuun tietojoukkoon [SW04, JAM06, s. 70]. Tavallisimpia operaatioita ovat arkiston kokoelmarakenteen selvittäminen sekä kokoelman tai yksittäisen viitteen metatie-

don hakupyynnöt (kuva 4). Käytännöllinen työkalu protokollan toiminnallisuuden tutkimiseen on selaimen kautta käytettävä OAI Repository Explorer¹⁹.

Merkittävimmät haasteet OAI-protokollan käyttöön liittyvät lähinnä tiedonhallintaan ja suurten tietomassojen käsittelyyn [JAM06, s. 84]. Yksi tällainen on toiseisen tiedon ongelma: arkistojen määrän lisääntyessä ja metatiedon haravoinnin yleistyessä päällekkäisiä viitetietoja kumuloituu yhä useampaan arkistoon [WN03]. Verkkoaineiston yksilöimiseen [Hak02] ja versiointiin liittyvät ongelmat hankaloittavat tällaisten duplikaattien poistamista. Pahimmassa tapauksessa arkistoihin kertyy kymmenittäin samaa artikkelia kuvaavia viitteitä, mikä edelleen huonontaa käytettävyyttä ja kuluttaa turhaan järjestelmän resursseja.

Protokollan määrittelyyn on myös esitetty parannuksia. Esimerkiksi sen suhtautuminen Dublin Coren tarkentimiin jää määrittelyssä hieman avoimeksi, minkä vuoksi osa toteutuksista jättää niiden sisällön käsittelemättä. Moniulotteisen metatiedon tuelle on myös selkeä tarve. Tällä hetkellä MHP toimii ainoastaan tiedonsiirtovälineenä, eikä sen kautta ole mahdollista tehdä esim. asiasanahakuja arkistoihin. Tällainen ominaisuus olisi kuitenkin tarpeen tehokkaiden ja nykyistä tarkemmin rajattujen viitetietojen keräämistä silmälläpitäen. Haasteena on myös käyttöoikeustietojen sisällyttäminen haravoituun metatietoon sekä viitteiden synkronointi alkuperäisen tiedon päivittyessä [VdSNL04, WN03]. Myös vertaisverkoissa toimivia toteutuksia on esitelty suorituskyvyn parantamiseksi [ANS02].

Metatiedon ja arkistoitavan materiaalin muuttuminen yhä monimuotoisemmaksi on tuonut uusia haasteita myös julkaisuarkistojen väliseen tiedonsiirtoon. OAI-työryhmä aloittikin lokakuussa 2006 uuden ORE-protokollan (Object Re-use and Exchange) määrittelyn. Tavoitteena on kehittää tiedon haravointimahdollisuuksia siten, että kuvailutiedon lisäksi voidaan jakaa myös niihin liitettyjä tiedostoja ja palveluita [LVdS07].

¹⁹<http://re.cs.uct.ac.za/>



Kuva 4: OAI-MHP protokollan tiedonsiirto.

3.5 Pysyvät tunnisteet verkossa

Tavallisesti johonkin Internetissä sijaitsevaan kohteeseen viitataan URL:n (Universal Resource Locator) avulla, joka määrittelee objektin absoluuttisen sijainnin verkossa. Tietolähteen vaihtaessa paikkaa jäljelle jää toimimaton linkki, jonka päivittäminen useaan paikkaan voi olla hankalaa tai mahdotonta. Ongelman laajuudesta kertoo se, että noin kolmasosa verkossa ilmestyneistä artikkeleista ei ole saatavilla enää kymmenen vuoden kuluttua niiden julkaisemisesta [MCNB05]. Tämän vuoksi pysyvien tunnisteiden luomisesta ja ylläpitämisestä on tullut tärkeä osa tiedon verkkojulkaisemista ja pitkäaikaistallennusta. Niiden avulla toivotaan myös ratkaistavan verkkomateriaalin identifiointiin liittyviä ongelmia [Sch03].

Tilannetta korjaamaan on suunniteltu järjestelmiä, jotka tarjoavat muuttumattoman osoitteen halutulle kohteelle Internetissä. Niiden toimintaidea on yksinkertainen uudelleenohjaus; kohteen absoluuttisen osoitteen käyttö korvataan helpommassa muodossa olevalla palvelunimellä. Jokainen palvelunimi muodostaa viitteen keskitettyyn hakemistoon, joka sisältää dokumentin absoluuttisen osoitteen verkossa. Käyttäjän osoittaessa tällaista pysyvän tunnisteiden muodostamaa linkkiä haetaan hakemistosta kyseiselle resurssille ajan tasalla oleva absoluuttinen osoite. Näin tie-

tolähteen vaihtaessa paikkaansa ainoastaan tunnisteita ylläpitävä taho joutuu päivittämään osoiterekisteriään ja loppukäyttäjää pääsee resurssiin edelleen käsiksi saman tunnisteiden kautta. Luonnollisesti ratkaisu toimii kuitenkin vain niin pitkään kuin kyseisen palvelun ylläpitäjä on olemassa. Tunnettuja pysyvien tunnisteiden luomiseen käytettyjä standardeja ovat mm. CNRI:n Handle, sen päälle rakennettu DOI (Digital Object Identifier) sekä PURL.

Kaupallisia järjestelmiä korvaamaan on hiljattain syntynyt kansalliskirjastojen hallinnoima URN-tunnistejärjestelmä. Sen toiminta rakentuu jo olemassa olevien tunnusjärjestelmien hyväksikäyttöön palvelunimeä rakennettaessa, mikä mahdollistaa julkaisun identifiointin jo pelkän osoitteen perusteella (taulukko 1). Mahdollisia tunnusjärjestelmiä on useita, mutta julkaisujen kohdalla erityisesti ISSN- ja ISBN-tunnukset ovat käyttökelpoisia. Jos viitattava kohde ei kuulu minkään sopivan tunnusjärjestelmän piiriin, kansalliskirjastot voivat jakaa erityisiä NBN (National Bibliographic Number) -numeroita [Sch03, Kes07].

Paveluosoite	Nimiavaruus	Tunnisteosa
http://hdl.handle.net/	2455/	19717
http://dx.doi.org/	10.1080/	07853890600630567
http://urn.fi/URN:	ISBN:	978-951-98479-4-8
http://urn.fi/URN:	ISSN:	278-931-93479-4-8
http://urn.fi/URN:	NBN:fi-	fe20061586

Taulukko 1: Pysyvät tunnisteet ja niiden rakentuminen.

URN-tunnisteet tarjoavat muita monipuolisemmat nimiavaruudet, joissa on selkeästi näkyvissä kohteen identifiointi tunniste. Tällä tavoin dokumentti on yksilöitävissä ja duplikaatit tunnistettavissa vaikka ne sijaitisivat eri paikassa. Lisäksi kansallisella tasolla ylläpidetyn palvelun jatkuvuus on erinomainen pitkälläkin aikavälillä [Kes07]. URN-tunnisteiden käyttöönotto selkeyttää verkkomateriaalin käyttöä ja identifiointia ja parhaimmillaan tuo myös ratkaisun toisteisen tiedon ongelmaan.

Vaikka verkkodokumentti olisi yksilöitävissä, ei käyttäjä voi tietää siihen kohdistuneista muutoksista. Ongelma on ollut esillä erityisesti tutkimusaineiston verkkotallennuksen kohdalla. Viimeisin innovaatio onkin muodostaa tunnistenimi käyttäen

hyväksi materiaalista laskettua hajautuskoodia (hash). Jos alkuperäinen data muuttuu, tuottaa sille laskettu hajautuskoodi eri tuloksen kuin ennen, eikä siitä muodostettu tunniste enää osoita samaan paikkaan kuin ennen [AK07].

Julkaisuarkistoa rakennettaessa tiedon saavutettavuuteen on kiinnitettävä erityistä huomiota; dokumenttien tulee olla tavoitettavissa samalla tunnisteella vaikka tekninen alusta tai sitä ylläpitävä taho vaihtuisi. Tämän vuoksi arkistoihin tallennetuille dokumenteille luodaankin syöttöprosessin yhteydessä pysyvät tunnisteet käyttäen hyväksi joko olemassa olevia tai arkistosovellusten omia tunnistejärjestelmiä [SBB03]. Niiden avulla verkkodokumentteihin viittaaminen on helppoa ja yksinkertaisempina kuin todelliset verkko-osoitteet ne toimivat myös paremmin painetun tekstin seassa, kuten artikkeleissa tai julkaisuluetteloissa.

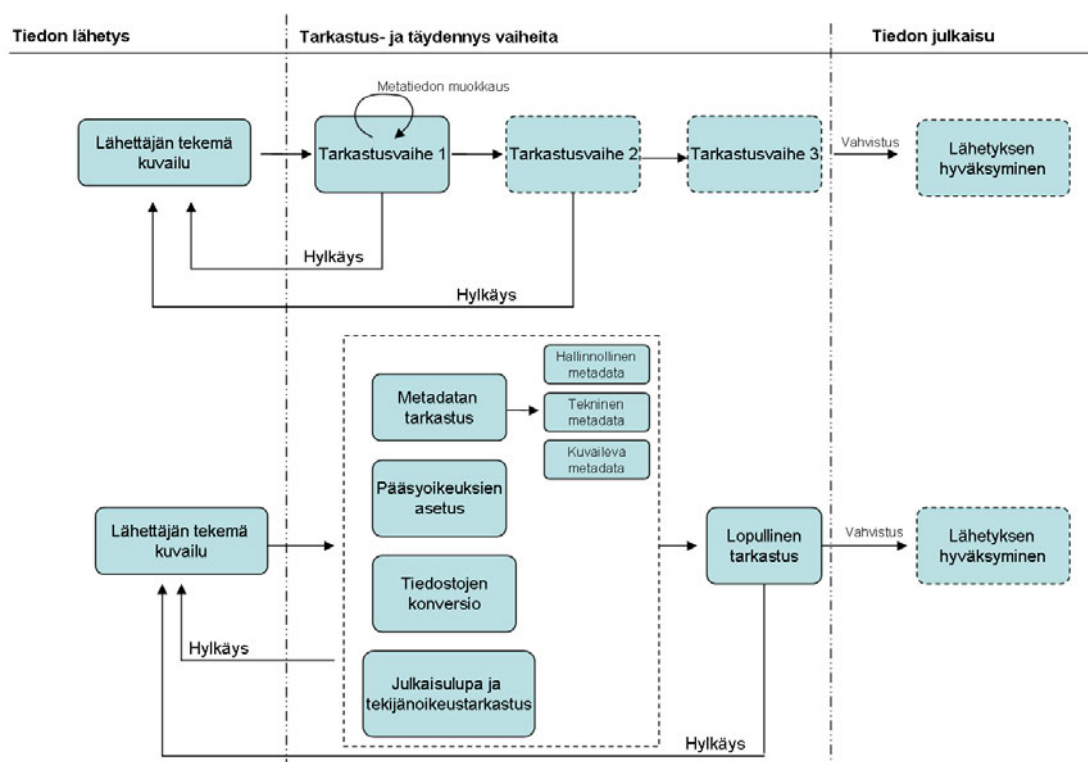
3.6 Työnkulun hallinta

Työnkululla tarkoitetaan niitä osavaiheita, jotka arkistoon syötetty dokumentti käy lävitse ennen päätymistään käyttäjien nähtäville. Julkaisuarkistojen on kyettävä tarjoamaan työkaluja erilaisten työnkulkujen laadintaan, joiden avulla materiaalin syöttö ja tarkastus voidaan toteuttaa joustavasti ja tarvittaessa hajautetuissa osavaiheissa eri henkilöiden toimesta [JAM06, s. 85]. Ohjelmistojen tarjoamat ominaisuudet eivät aina kuitenkaan riitä, minkä vuoksi toimivan syöttöprosessin kehittäminen vaatii tiivistä yhteistyötä järjestelmän käyttäjien sekä teknisen ylläpidon kanssa.

Tavallisesti materiaalin tyyppi sekä käytettävissä olevat henkilöstöresurssit määrittelevät aineiston käsittelyvaiheiden määrän ja järjestyksen. Usein päivittyvä tai muuten aikakriittinen materiaali, kuten oppimateriaali, vaatii suoraviivaisen ja välikädetömän syöttöprosessin. Sen sijaan julkaistut artikkelit, jotka kuvaillaan täsmällisesti, hyötyvät moniportaisesta tarkastusprosessista eikä pientä viivettä niiden julkaisemisessa useinkaan koeta ongelmaksi.

Tyypillisiä työnkulun osavaiheita ovat tiedon syöttö, metatiedon tarkastus sekä julkaisun hyväksyminen. Syöttövaiheessa tiedon lähettäjä kuvailee lähettämänsä materiaalin verkkolomakkeen avulla. Tarkastusvaiheessa lähetyksen tiedot täydennetään ja saatetaan vastaamaan sovittuja standardeja. Hyväksymisvaiheessa tarkastetul-

le lähetykselle puolestaan annetaan lopullinen julkaisulupa, minkä jälkeen se tulee kaikkien nähtäväksi arkistoon. Näiden lisäksi tarpeellisia osavaiheita voivat olla mm. tekijänoikeuksien tarkastus tai materiaalin konversio. Kullekin osavaiheelle voidaan määritellä omat vastuuhenkilöt aineiston tyypistä tai erityisvaatimuksista riippuen. Toisinaan joitain vaiheita voi olla myös tarpeen hajauttaa useaan osaan. Esimerkiksi moniulotteisen kuvailutiedon eri osien tarkastus voidaan tällöin jakaa useammalle vastuuhenkilölle. (kuva 5).



Kuva 5: Kaksi erilaista työnkulkua; yllä lineaarinen malli, alla rinnakkaisia vaiheita sisältävä malli.

4 Julkaisuarhivistot opetuksen ja oppimisen tukena

Perinteinen opetuksen malli rakentuu ajatukselle, että opettaja pitää hallussaan oppimisen arvoista tietoa, joka tulisi saada jalkautettua opiskelijoiden keskuuteen. Tietoyhteiskunnan kehittymisen ja digitaalisen vallankumouksen myötä oppimisen ja opettamisen käsitteet sekä tiedon tuottamisen malli ovat kuitenkin muuttuneet

perustavalla tavalla [Roe01, MM95]. Yhä monimutkaistuvien tutkimuskohteiden äärellä myös opettajan ja oppijan roolit hämärtyvät. Opettajan rooli onkin muuttumassa tiedon jakajasta tiedon arvioinnin ja analyysin – informaatiolukutaidon – opettamiseen [MM95, HLI99]. Oppimiskäsityksessä tapahtuneiden muutosten myötä myös sitä tukemaan suunniteltujen oppimisympäristöjen on kehityttävä.

Tulevaisuuden oppimisympäristöille voidaan ajatella olevan ominaista opiskelijakeskeisyys, vuorovaikutteisuus ja dynaamisuus sekä tosielämän ongelmien ratkominen ryhmätyönä. Tieteellinen tieto ja sen myötä informaatiolukutaidon korostuminen ovat myös keskeisessä roolissa [DC07, TM98]. Tästä johtuen näyttäisikin siltä, että digitaalisilla tietovarastoilla, kuten julkaisuarkistoilla, tulee olemaan merkittävä rooli oppimisympäristöjen kehityksessä ja oppimisprosessin tukemisessa [DC07].

4.1 Oppimisen malleja

Pedagogiikan, oppimisteorioiden ja teknologian vuorovaikutus on mahdollistanut uusia opetuksen ja oppimisen malleja [DHM04]. Osana tätä merkittävää muutosta esiin ovat nousseet mm. tutkivan ja sulautuvan oppimisen käsitteet. Seuraavaksi esitellään lyhyesti em. malleja sekä pohditaan mahdollisuuksia hyödyntää julkaisuarkistoja niiden tarjoamassa oppimisen viitekehyksessä.

Sulautuva oppiminen

Sulautuvalla oppimisella (Blended learning) tarkoitetaan oppimisprosessia, jossa oppimisympäristöt ja niiden sisältämät erilaiset opetusmuodot ja välineet yhdistyvät uudenlaiseksi, oppimista edesauttavaksi kokonaisuudeksi. Näin ollen opetus voi koostua tavanomaisesta opettajajohtoisesta luokkaopetuksesta, asiantuntijaluennoista, itseopiskelusta sekä itseorganisoidusta tai verkkoympäristöissä tapahtuvasta opiskelusta. Formaalin opiskelun ohella oppimiseen voi liittyä ei-formaalia toimintaa kuten työharjoittelua ja mentorointia, jotka kukin edistävät tiedonhankintaa ja uuden oppimista [DHM04]. Usein käsitteeseen liitetään vahvasti tieto- ja viestintätekniikan tuomat mahdollisuudet, joskin mallia voi toteuttaa ilman niitäkin. Erona jo pitkään toteutettuun monimuoto-opiskeluun lienee tietoinen pyrkimys oppimistilanteen tar-

kasteluun kokonaisvaltaisemmin kuin yksittäisten välineiden tai oppimistilanteiden tasolla [LJR05]. Mallin voitaneen ajatella pyrkivän yhdistämään lähiopetuksen tehokkuuden ja sosiaaliset mahdollisuudet uuden teknologian mahdollistamaan aktiivisempaan, opiskelijälähtöisempään, oppimistapaan [DHM04]. Opetusmuotojen runsaus tuo lisäksi joustavuutta ja valinnanmahdollisuuksia kaikille osapuolille.

Ongelmakeskeinen oppiminen

Ongelmakeskeinen oppiminen (Problem-Based Learning) lähtee liikkeelle siitä, että oppiminen on aina tilannesidonnaista. Pureutumalla opetustilanteessa tosielämän ongelmatapauksiin opiskelu koetaan mielekkäämpänä ja opittu aines on helpompi omaksua. Mm. Helsingin yliopistossa lääketieteen pedagogiikka onkin jo pitkään rakentunut ongelmalähtöisen oppimisen periaatteelle. Käytännössä opetus toteutetaan usein paneutumalla pienryhmissä todellisiin potilastapauksiin ja rakentamalla niiden valossa ymmärrystä tutkittavasta asiasta tai ilmiöstä laajemmin [CK04].

Opetustilanteessa osallistujilla on selkeät roolit; puheenjohtaja ohjaa keskustelua, sihteeri kirjaa tehtyjä havaintoja ylös ja muut opiskelijat osallistuvat keskusteluun. Opettaja suuntaa ryhmän toimintaa oikeaan suuntaan ja toisinaan käytössä voi olla myös asiantuntija koordinaattorin roolissa. Oppimisprosessi jakautuu seitsemään vaiheeseen, joita ovat käsitteiden selventäminen, lähtökohtien määrittely, aivoriihi, selitysmallin rakentaminen, oppimistavoitteiden muotoilu, itseopiskelu sekä opitun tiedon analysointi ja ongelman ratkaiseminen. Osavaiheiden aikana yhdessä kootut ideat, ratkaisuesitykset sekä tavoitteet kirjataan ylös ja itseopiskeluvaiheen jälkeen niihin palataan tiedonhankinnan viisastamana. Lopuksi kootaan ongelmalle selitys sekä arvioidaan kuinka hyvin sen ratkaisussa onnistuttiin [HLL99].

Tutkiva oppiminen

Tutkiva oppiminen on oppimisen malli (kuva 6), jossa oppiminen ei perustu opettajan tai oppikirjan jakaman faktatiedon opiskeluun, vaan oppijan omaehtoiseen ja ohjattuun tiedonrakenteluun osana yhteisöä. Keskeistä on myös, ettei uutta tietoa omaksuta suoraan aikaisempiin ajatusmalleihin, vaan arvioinnin ja uuden tiedon

etsinnän tuloksena syntyneistä oivalluksista pyritään rakentamaan uusia laajempia kokonaisuuksia. Tarkoituksena on luoda oppimisprosessi, joka jäljittelee asiantuntijaorganisaatioille tyypillisiä tiedonrakentamisen käytäntöjä. Tutkivan oppimisen kulmakiviä ovat ilmiöiden entistä laajempi ymmärtäminen ja selitysten tuottaminen, ongelmalähtöinen tiedonhankinta, ennakkokäsitysten reflektointi uuteen tietoon sekä yhteisöllinen tiedonrakentelu ja jaettu asiantuntijuus [HLI99].

Tutkivan oppimisen mallissa oppija ottaa vastuun oman oppimisprosessin onnistumisesta työstämällä ja käsitteellistämällä sitä aktiivisesti ja luoden näin perustaa korkeamman asteen tiedonkäsittelytaitojen kehittymiselle. Oppijaa rohkaistaan luomaan selityksiä ja syventäviä kysymyksiä tutkimuksen kohteena oleville ilmiöille sykleittäin etenevissä jaksoissa, joista jokainen syventää asian käsittelyä ja tuottaa myös joukon lisäkysymyksiä.

Tutkittava ongelma ankkuroidaan ensin johonkin todellisen maailman ongelmaan ja oppijoiden omiin kokemuksiin, minkä jälkeen määritellään ongelmat. Niiden rooli on mallissa keskeinen, sillä ne ohjaavat tutkimuksen kulkua jatkossa ja parhaimmillaan motivoivat vanhan ja uuden yhteensovittamisessa. Seuraavaksi kuvaillaan ja selitetään ilmiö nykyisten taustatietojen varassa ja luodaan niistä työskentelyteorioita kirjoittamalla ja keskustelemalla. Syventävän tiedon hakeminen edeltää ja seuraa kriittisen arvioimisen vaihetta, jonka aikana pyritään punnitsemaan esitietojen pohjalta luotuja käsityksiä sekä ratkomaan tunnistettuja ongelmia. Hankitun tiedon pohjalta kysymyksenasettelua ja teorioita voidaan tarkentaa ja prosessi voidaan haluttaessa aloittaa uudestaan. Kaikissa vaiheissa yhteistoiminnallisuudella ja keskustelulla on suuri rooli; sosiaalinen vuorovaikutus näyttäisikin toimivan välineenä synnytettyjen ideoiden testaamisessa. Taustalla on ajatus yksilön roolin muuttumisesta tietoyhteiskunnassa, missä osaaminen näyttäytyy toisiaan täydentävien asiantuntijoiden verkostona [HLI99].

4.2 Julkaisuarkisto osana oppimisprosessia

Edellä esitettyjen oppimisen mallien ajatellaan soveltuvan hyvin tietoyhteiskunnan oppimisen tarpeisiin. Malleja yhdistäviä tekijöitä ovat tutkimustiedon hyödyntä-



Kuvio 1.
Tutkivan oppimisen prosessi

Kuva 6: Tutkivan oppimisen malli [HLI99, s. 13].

minen ja uuden tiedon etsiminen osana ongelmanratkaisua sekä tutkimuskysymysten kriittinen arviointi hankitun tiedon valossa. Lisäksi keskeiselle sijalle nousevat ryhmätyötaidot ja oman osaamisen jakaminen yhdessä muiden kanssa. Sulautetun oppimisen viitekehyksessä opetuksessa pyritään hyödyntämään erilaisia väyliä, tekniikoita ja välineitä mahdollisimman moniulotteisen oppimiskokemuksen toteuttamiseksi. Julkaisuarkistot voivat teknisinä alustoina olla mahdollistamassa tällaisen oppimisprosessin toteutumista, jossa kaikki edellämainitut elementit ovat vahvasti läsnä. Olemassaolevista oppimisympäristöistä poiketen ne tarjoavat valmiin materiaalivaraston, rajapinnan tieteen tekemisen todellisuuteen sekä osallisuuden asiantuntijaverkostoon (kuva 7).

Ominaisuuksiensa puolesta arkistot tarjoavat valmiita välineitä tiedon säilömiseen, tehokkaiden hakujen tekemiseen sekä omien tuotosten julkaisuun ja vertaisarviointiin. Tähän mennessä ne ovatkin löytäneet paikkansa opetuskäytössä lähinnä tietovarastoina, joihin erillisistä oppimisympäristöistä viitataan [Roe01]. Uusien oppimisen mallien lähestyessä asiantuntijatyön prosesseja ne voidaan valjastaa myös oppimista edistävään käyttöön. Haasteena onkin kehittää niistä järjestelmiä, jotka tukevat

aidosti oppimisprosessia sen eri vaiheissa eivätkä toimi pelkästään jakelukanavana materiaalille.

Tässä erityisesti oppijoiden verkottuneisuus ja sosiaalinen media voivat olla avaamassa uusia mahdollisuuksia, jotka ovat aiemmin saaneet vain vähän huomiota opetuskäytössä [Kru06, Ers06]. Pienellä lisätyöllä arkistosovelluksiin on helposti liitettävissä sosiaalista vuorovaikutusta ja prosessikirjoittamista tukevia lisäpalveluja kuten pikaviestimiä, wikejä ja blogeja. Niiden kautta yksisuuntainen julkaisukanava voidaan muuttaa vuorovaikutteiseksi asiantuntijoiden ja opiskelijoiden foorumiksi. Tällaisen järjestelmän ympärille olisi mahdollista rakentaa todellista yhteisöllistä oppimista, jossa tieteen tekeminen ja oppiminen integroituvat saumattomasti. Perinteisiä verkko-oppimisalustoja riippumattomampina ne tarjoavat mahdollisuuden globaaliin verkosto-oppimiseen, missä useista eri taustayhteisöistä tulevat osallistujat kokoontuvat oppimaan formaalisti, epäformaalisti tai ammattinsa puolesta [MM95].

Arkistojen toimiessa luonnollisena risteyspaikkana asiantuntijoille ja opiskelijoille ne voivat samalla tarjota heidän käyttöönsä laajan valikoiman erilaisia lisäpalveluja elinikäisen oppimisen ja tiedonrakentelun tueksi. Tarjoamalla oppija- tai työryhmäkohtaisia profiileja ja dokumenttikoreja järjestelmä voi toimia eräänlaisena oppimisen tiekarttana tai opintorekisterinä pitkälläkin aikavälillä [Roe01]. Jo nyt käytössä olevat tutkijaportfoliot vapauttavat tutkimuksen tekijät julkaisuluetteloiden ylläpitämiseltä omilla kotisivuillaan sekä tarjoavat heille yksityistä työskentely- ja tallennustilaa [GA08]. Opiskelijat voisivat hyötyä samanlaisesta palvelusta. Lisäksi semanttisen webin ja tiedonlouhinnan työkalujen avulla arkiston valtavasta tietomassasta voitaneen tulevaisuudessa profiloida käyttäjäkohtaisia palveluja opintopiirien ja asiantuntijaverkostojen luomiseen tai oman tuotannon kanssa yhtenevien julkaisujen etsimiseen [SBCH06].

Esimerkkitapaus

Julkaisuarkiston ympärille rakennettu oppimisprosessi voi toteutua esimerkiksi seuraavalla tavalla. Opetuksen alkaessa konteksti tai ongelmanasettelu poimitaan arkistoon tallennetusta tutkimusdatasta tai artikkelista. Sen pohjalta asetetaan tut-

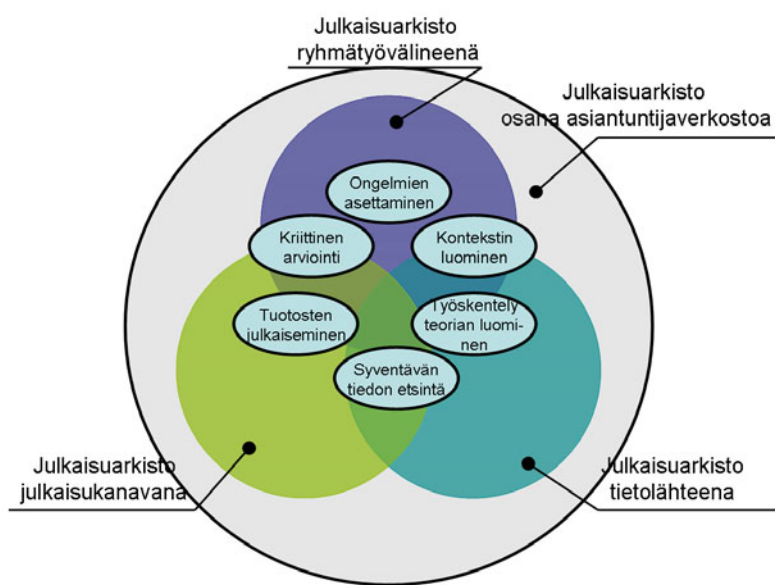
kimusongelmat sekä määrittellään työskentelyteoriat. Tämä voidaan toteuttaa sekä ryhmätapaamisissa että verkon tarjoaman sosiaalisen median välityksellä, kuitenkin ohjaavan opettajan kontrollissa. Arkiston tarjoama työryhmäkohtainen wiki voi tarjota sopivan alustan ongelman määrittelylle ja työstämiselle. Työskentelyteorioiden ja pohdiskelujen julkaiseminen mahdollistaa tuotosten opponoinnin ja omien teorioiden kriittisen arvioinnin. Myös koko oppimisprosessin kulkua ja ohjautumista voidaan tällöin vertailla, mikä mahdollistaa tehokkaiden oppimistapojen oppimisen.

Jaettu asiantuntijuus toteutuu luonnollisesti sekä lähiopetustilanteissa että verkossa. Verkkokeskusteluissa ja pikaviestimien kautta käyty aivoriihi mahdollistaisi lähitapaamisten erikoistumisen opettajan prosessia suuntaavaan palautteeseen ja asiantuntijavieraiden hyödyntämiseen osana niitä. Julkaisuarkiston tarjoaman asiantuntijaverkoston myötä huippuosaaajia voi olla mahdollista hyödyntää opetuksessa entistä helpommin kommentaattoreina verkkokeskusteluissa. Kenties opiskelijat hyötyisivät myös samaan aikaan muiden ongelmien parissa puurtavien opiskelijaryhmien kanssa keskustelualueilla käydystä ajatustenvaihdosta.

Sekä lähiopetuksessa että verkossa saadun palautteen perusteella ongelmiin etsitään lisäselvennystä. Tällöin julkaisuarkistot voivat toimia yhtenä merkittävänä tietolähteenä oppikirjojen ja muun materiaalin rinnalla. Tässä vaiheessa myös tiedon hakemisen opettelu ja sen relevanssin arviointi, informaatiolukutaito, korostuu entisestään; löydettyä tietoa on osattava arvioida ja hyödyntää oikein. Itseopiskelun ja uusien havaintojen pohjalta keskustelua jatketaan ja ideoita kirjataan ylös sekä verkon että lähitapaamisten avulla. Myös aihepiirin kannalta relevantti tutustumiskäynti sekä muut ei-formaalit oppimisen muodot voivat olla mukana syventämisessä oppimista. Ideaalitilanteessa koko oppimissykli voitaisiin integroida osaksi harjoitustyön tekemistä tai todellista tutkimusprojektia, jolloin siihen saataisiin myös käytännön ulottuvuus. Tutkimusryhmien hyödyntäessä julkaisuarkistoa omassa toiminnassaan, myös opiskelijoilla olisi mahdollisuus seurata niiden työn etenemistä ja oppia siitä.

Syklin lopuksi työstetään uusi, alkuperäisestä muokattu työskentelyteoria, joka myös julkaistaan arkistossa. Tutkimusprosessin tuotosten julkaisemisen voidaankin katsoa olevan tärkeä arvioinnin muoto, joka pakottaa opiskelijan tarkastelemaan tekstiään

entistä syvällisemmin myös lukijan näkökulmasta. Lisäksi aiemmin tuotettua ja arkistoitua materiaalia voidaan käyttää muiden oppijoiden toimesta edelleen osana heidän omaa oppimisprosessiaan [HLI99, s. 54]. Tämän jälkeen tutkimussykli voidaan haluttaessa aloittaa uudestaan määrittelemällä tutkimuksen konteksti uudesta työskentelyteoriasta käsin (kuva 7).



Kuva 7: Julkaisuarkisto oppimisprosessin osana.

5 Meilahden kampuksen avoin julkaisuarkisto

Helsingin yliopiston lääke- ja hammaslääketieteen opetus- ja tutkimustoiminta sekä sairaanhoito on keskittynyt Meilahden kampukselle. Päivittäin se on 2000 opiskelijan ja 1500 opettajan työpaikka. Alueella toimii myös 1200 tutkijalle työtilat tarjoava tutkimus- ja opetuskeskus Biomedicum.

Meilahden kampuksen avoimen julkaisuarkiston kehittämisprojekti alkoi Terveystieteiden keskuskirjaston (Terkko) toimesta vuoden 2005 lopussa olemassa olevien järjestelmien kartoituksella ja testiympäristön rakentamisella. Puolentoista vuoden kehitystyön tuloksena syntyi Terkko Document Space (TDS) -palvelu. Virallisesti

se julkistettiin IT Medi -seminaarissa Biomedicumissa toukokuussa 2007²⁰. Arkistotullaan ottamaan laajempaan käyttöön asteittain vuoden 2008 aikana kehitystyön edelleen jatkuessa.

5.1 Käyttötarkoitus ja motivaatio

Julkaisuarkiston pääasiallinen tarkoitus on säilöä kampuksella tuotettua tieteellistä kokotekstiaineistoa sekä muuta lääketieteeseen liittyvää historiallista ja ajankohtaista materiaalia. Tällaista ovat mm. tutkimusartikkelit, seminaariesitelmät ja historialliset väitöskirjat. Palveluun tallennetaan myös julkaistujen artikkelien ja dokumenttien metadatasia, jota haetaan automaattisesti kotimaisista ja kansainvälisistä tietokannoista. Metatieto toimii pääasiassa raakadatan paikallisesti syötetyn aineiston kuvailua varten.

Myös lääketieteellisen tiedekunnan oppimateriaalin jakamiseen käytetty digitaalinen kurssikirjasto (DiKK) tullaan siirtämään osaksi kampuksen julkaisuarkistoympäristöä. Jo lähes kymmenen vuoden ikään ehtinyt palvelu kerää vuosittain yli 30 000 käyttäjäistuntoa. Se sisältää noin 700 opetuskäyttöön suunniteltua dokumenttia, pääasiassa luentokalvoja. Alkuun molemmat palvelut toimivat rinnakkain omina kokonaisuuksinaan, mutta ajan mittaan ne on tarkoitus siirtää yhteen järjestelmään. Tutkimus- ja opetusmateriaalivarastojen yhdistäminen viestii näiden kahden tiedonrakentelun kentän voimakkaasta integraatiosta; tutkiva oppiminen ja kampuksella tehty tieteellinen tutkimus lähentyvät toisiaan.

Lääketieteisiin keskittyneestä julkaisuarkistosta ovat kiinnostuneet Meilahden kampuksen lisäksi myös joukko yliopiston sektoritutkimuslaitoksia sekä muita terveystieteiden parissa toimivia tahoja. Näistä Kansanterveyslaitos, Stakes, Kansaneläkelaitos ja UKK -instituutti ovat jo ilmoittaneet halukkuutensa arkistoida tuottamaansa materiaalia palveluun. Jatkossa järjestelmään tultaneen arkistoimaan myös alan kotimaisia lehtiä. Kampuksen tarpeista liikkeelle lähtenyt palvelu onkin nopeasti kehittymässä Suomen johtavaksi terveystieteiden julkaisuarkistoksi.

²⁰http://www.ltdk.helsinki.fi/tuke/tapahtumat/IT-Medi_07.htm

5.2 Sovelluksen valinta

Toimivan palvelun rakentamiseksi järjestelmän huolellinen valinta katsottiin ensiarvoisen tärkeäksi. Puolentoista vuoden hankkeen kokonaisajasta noin kolmasosa käytettiin sopivan arkistosovelluksen etsimiseen ja testaamiseen.

Projektin alussa laadittiin karkea vaatimusmäärittely ominaisuuksista, joita järjestelmällä haluttiin olevan. Sen pohjaksi otettiin opetusteknologiakeskuksen toimesta jo aiemmin tehty vaatimusmäärittely (Liite 1). Lisäksi pohdittiin järjestelmän ylläpitoon ja kehitykseen liittyviä vaatimuksia, erityisesti Terkon resursseja jatkokehittää ja räätälöidä valittua ohjelmistoa kampuksen erikoistarpeita varten. Järjestelmältä peräänkuulutettiin myös mahdollisuutta tallennetun tiedon siirtämiseen ulos josain standardissa muodossa. Sen avulla arkisto voi jatkaa elämäänsä tulevaisuudessa myös jollain muulla alustalla. Näiden pohjalta tehtyjä havaintoja peilattiin vielä muualla tehtyihin arviointidokumentteihin [Wyl06] sekä eri järjestöjen julkaisemiin suosituksiin [Cro02, BW05].

Alkukartoituksen jälkeen ohjelmistoista CDS Invenio ja DSpace valittiin syvällisempää kokeilua varten. Molemmilla näistä oli omat vahvuutensa ja ominaispiirteensä (Liite 1). CDS Invenio vakuutti ominaisuuksillaan ja suoristuskäytöllään. Monet lupaavista toiminnoista olivat kuitenkin pahasti keskeneräisiä eikä dokumentointia juurikaan ollut saatavilla, mikä johtunee ohjelmiston varhaisesta kehitysvaiheesta. Järjestelmän asennus ja käyttöönotto osoittautui myös poikkeuksellisen vaativaksi, eikä pieni käyttäjäyhteisö tarjonnut riittävästi näyttöä sovelluksen käyttömahdollisuuksista. Invenion sekalaiset tekniset ratkaisut (Perl, Python) ja niiden tuntemuksen puute kampuksella vaikeuttivat sovelluksen räätälöintiä ja jatkokehitystä. Myös muualla tehdyt kokeilut tukivat vaikutelmaa järjestelmän keskeneräisyydestä [Wyl06, s. 13].

DSpace osoittautui tarjolla olevista järjestelmistä kypsimmäksi ja pisimmälle viedyksi kokonaisuudeksi. Sitä käytetään tieteellisen työn arkistointiin sadoissa yliopistoissa ympäri maailmaa ja Helsingin yliopistossakin mm. Viikin kampuksella ja Kansalliskirjastolla on siitä aiempaa kokemusta. Teknisten ratkaisujensa johdosta (Java) kampuksen resurssit sen jatkokehittämiseen ovat hyvät. Lisäksi kampusarkis-

tojen yhteiskäyttö sekä jatkokehitys lienevät helpompaa teknisten alustojen ollessa samanlaiset. Terävin kritiikki DSpaceen kohdalla osuu ohjelmiston skaalautuvuuteen suurilla tietomäärillä, minkä on raportoitu aiheuttavan suorituskykyongelmia [Wyl06, s. 16]. Järjestelmän kehitysnäkymät ovat kuitenkin positiiviset ja tuoreen tiekartan mukaan ohjelmistoa pyritään jatkossa kehittämään nimenomaan tältä osin, mm. arkkitehtuuri- ja tietokantauudistuksin [Tan07, Ock07]. Näin ollen Meilahden kampuksen julkaisuarkisto-ohjelmistoksi valittiin DSpace.

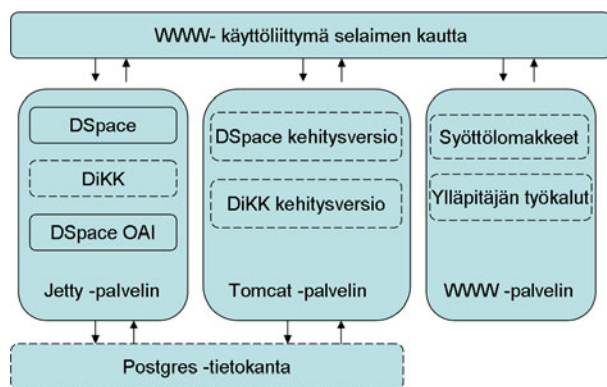
5.3 Järjestelmän tekninen kuvaus

Terveystieteiden keskuskirjasto vastaa Meilahden kampuksen julkaisuarkiston kehityksen lisäksi sen teknisestä ylläpidosta ja toimivuudesta. Palvelu on asennettu paikallisesti ylläpidettävälle palvelimelle kirjaston palvelinhuoneeseen. Materiaalin ja käyttäjämäärien lisääntyessä harkittaneen myös järjestelmän siirtämistä tietotekniikkaosaston tarjoamalle virtuaalipalvelimelle.

Tekninen alusta

Palvelimen käyttöjärjestelmänä on Debian Linux. Sen päällä ajettava Apache WWW-palvelin on yhdistetty mod/jk-rajapinnan kautta Tomcat- ja Jetty-sovelluspalvelimiin (servlet container), mikä mahdollistaa selaimen palvelupyyntöjen ohjaamisen suoraan niille. Tietokantana käytetään paikallisesti asennettua Postgresia.

Palvelimelle on asennettu kolme eri DSpace-instanssia, joihin käyttäjät ohjataan eri palveluosoitteiden kautta. Yksi näistä näkyy kaikille avoimena julkaisuarkistopalveluna (TDS) ja toinen toimii HY:n sisäverkkoon rajattuna tiedekunnan digitaalisena kurssikirjastona (DiKK). Kolmas toimii kehitysympäristönä erilaisia kokeiluja varten. Tuotantoversioita ajetaan Jetty-sovellusalustalla, mikä osoittautui suorituskykyisemmäksi kuin kehitysversiota pyörittävä Tomcat. Järjestely mahdollistaa kehitysympäristön vapaan käytön ilman käyttökatkoksia julkisissa palveluissa (kuva 8).



Kuva 8: Meilahden kampuksen julkaisuarkistoympäristö. Pisteviivan ympäröivät palvelut on rajattu saavutettavaksi vain yliopiston verkosta.

DSpace

DSpace lienee julkaisuarkistosovelluksista tunnetuin. Se sai alkunsa vuonna 2000 Hewlett-Packardin ja MIT laboratorioden yhteistyöprojektissa, mutta on kahdeksan vuoden aikana paisunut usean yliopiston ja aktiivisen harrastajayhteisön kehittämäksi avoimen lähdekoodin projektiksi. Ohjelmistoa kuvaillaan organisaation tutkimusmateriaalin tallennusta, pitkäaikaissäilytystä, indeksointia sekä jakelua silmälläpitäen kehitetyksi digitaaliseksi dokumenttivarastoksi [SBB03].

Ohjelmiston käyttöliittymän muodostaa WWW-sovellus, jota ajetaan jollain sovel-luspalvelimella. Java-pohjaisuutensa ansiosta DSpace on saatavissa kaikille käyttö-järjestelmille. Se hyödyntää useita kolmannen osapuolen ohjelmistoja, kuten Postgres-tietokantaa sekä CNRI:n Handle-muotoisia pysyviä tunnisteita. Metatietoa järjes-telmä käsittelee Dublin Core -muodossa.

DSpaceen arkistoitu tieto tallennetaan kahdessa osassa; metadata sekä varsinainen tiedosto (bitstream) säilötään erikseen. Yhdessä nämä muodostavat kokonaisuuden, josta käytetään nimeä viite tai tietue (content). Tietokantaa käytetään metatietojen säilömiseen yhdessä käyttäjätunnusten, kokoelmarakenteen sekä ohjelman muiden asetusten kanssa. Fyysiset tiedostot sen sijaan tallennetaan palvelimen tiedostojär-jestelmään, jonne viitteen metatiedoissa tarjotaan linkki [SBB03]. Lisäosien avulla DSpacen arkkitehtuuri sallii tietokantapohjaisen ratkaisun myös tiedostojen säilö-

miseen, mistä voi olla merkittävää hyötyä suuria tietomääriä siirrettäessä ²¹.

Ohjelmassa on kattavat oikeuksienhallintatyökalut, joiden avulla pääsy tiettyihin kokoelmiin tai dokumentteihin voidaan rajoittaa määrättyihin henkilöihin tai niiden muodostamiin ryhmiin. Käyttäjätunnistuksessa voidaan käyttää tunnusten sijaan myös sertifikaatteihin tai LDAP:iin perustuvaa tunnistusta. Muita kiinnostavia ominaisuuksia ovat mm. kokotekstin indeksointi, RSS-syötteet, Creative Commons-lisenssit sekä tuki OAI-protokollalle. Käyttäjäyhteisön tekemillä lisäosilla toiminnallisuutta voidaan laajentaa entisestään [SBB03].

Tunnettuja DSpace-tekniikalle rakennettuja julkaisuarkistoja ovat esim. MIT:n kurssikirjasto²² sekä Helsingin yliopiston kampusarkistot Viikissä ja Kumpulassa.

Manakin-käyttöliittymä

Julkaisuarkiston käyttöliittymältä vaaditaan paljon; sen tulee taipua palvelemaan monia eri käyttäjäryhmiä, tukea eri kieliversioita sekä mahdollistaa palvelun profilointi osaksi taustayhteisönsä muita verkkopalveluita. Loppukäyttäjälle kokemus palvelusta näyttäytyy aina sen käyttöliittymän kautta; hankalaksi koettua järjestelmää ei haluta käyttää.

DSpace-oma käyttöliittymä rakentuu jähmeästi JSP-sivujen joukkoon tuotetusta HTML-koodista, eikä mahdollista kovin pitkälle menevää ulkoasun yksilöintiä (Liite 2). Tätä varten Texasin yliopistossa onkin kehitetty lisäosa, jonka avulla arkistoon voidaan luoda kokoelma- tai jopa viitekohtaisia ulkoasuteemoja sekä yksilöllisiä näkymiä haluttuun osaan järjestelmän sisällöstä. Manakin-nimisen sovelluksen toiminta-ajatus on muodostaa DSpace-ohjelmiston koodista riippumaton XML-muotoinen esitys arkiston rakenteesta ja tietosisällöstä, minkä jälkeen siihen voidaan kohdistaa valikointia ja muokkaustoimenpiteitä. Lopullinen räätälöity esitys kustakin näytöstä tuotetaan XHTML-muodossa ja muotoillaan CSS-tyylien avulla [PGL05, PGM07].

TDS:n käyttöliittymä rakennettiin Manakinia hyväksikäyttäen. Sen avulla sivuston

²¹DSpace/SRB Integration Project: <https://libnet.ucsd.edu/nara/>

²²<https://dspace.mit.edu/>

ulkoasu saatiin helposti noudattamaan Lääketieteellisen tiedekunnan tyylimääritteilyjä. Myös monikielisen käyttöliittymän tuottaminen oli helppoa; kaikki käyttöliittymän tekstit ladataan dynaamisesti kielivalinnan mukaisesta kielitiedostosta. Sitä voidaan helposti myös täydentää palvelukohtaisten vaatimusten mukaisesti. Tällä hetkellä julkaisuarkistoa voi käyttää suomeksi, ruotsiksi ja englanniksi.

Käyttöliittymää rakennettaessa haluttiin välttää käyttäjän kokemus eksymisestä arkiston rakenteeseen. Navigoinnin helpottamiseksi käyttäjä näkee koko ajan sijaintinsa sivuston ylälaudassa olevassa palkissa. Palkin askeleet ovat linkkejä, joista voidaan siirtyä taaksepäin niitä vastaaville sivuille. Vasemmassa laidassa on aina näkyvillä pikalinkit tavallisimmin käytettyihin toimintoihin kuten kokoelma- ja viitelistaukseen, sekä koko arkistoon kohdistuva pikahakukenttä. Sinne on myös koottu linkkejä ohjesivustoille ja muihin HY:n julkaisuarkistoihin (Liite 3).

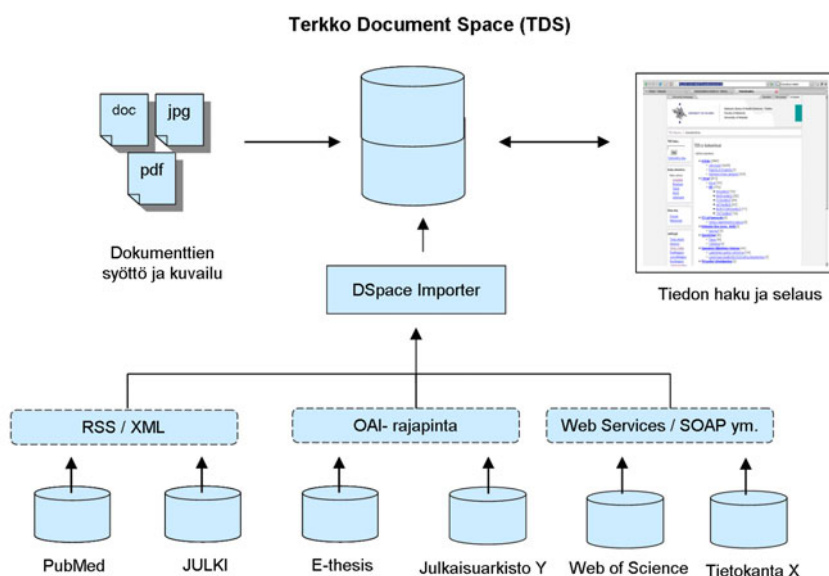
Yhteydet muihin tietokantoihin

Useat kustantajat mahdollistavat artikkeleidensa viitetietojen käyttämisen ilmaiseksi tai maksua vastaan. Tällöin ne ovat useimmiten haettavissa XML-muodossa palveluntarjoajan rajapinnan kautta. Julkaisuarkistokäytössä valmiin kuvailutiedon hyödyntäminen onkin houkuttelevaa, sillä se säästää aikaa ja vaivaa tallennusprosessin yhteydessä sekä tarjoaa täydelliset bibliografiset tiedot julkaisuista. Toisinaan saatavilla on myös kiinnostavaa lisätietoa, kuten viittausmääriä tai lehtien vaikutavuuskertoimia.

TDS hakee kahdesti päivässä uusimpia viitetietoja useasta eri lähteestä (kuva 9). Näitä ovat mm. PubMed- ja e-thesis-tietokannat. Lisäksi tietoa voidaan haluttaessa noutaa myös Thomsonin Web of Science ja Julki-tietokannoista. Hakuehtona voidaan käyttää Helsingin yliopistoa, merkittävien tutkijoiden nimiä tai viitteen yksilöivää tunnistetta, mikäli sellainen on tiedossa. Saatuja tuloksia käsitellään ja täydennetään tarpeen mukaan ennen niiden liittämistä osaksi paikallista arkistoa (kuva 9). Tämän jälkeen ne ovat hyödynnettävissä esim. raakadatana paikallisessa materiaalin syötössä. Open Access -ideologian mukaisesti myös TDS:n metatiedot ovat vapaasti halukkaiden käytettävissä. Tilastojen mukaan kampusarkiston viitteitä haravoidaan

tällä hetkellä kotimaisten palveluiden lisäksi ainakin muutamasta eurooppalaisesta yliopistosta käsin.

Metatiedon liian ahnaassa keräämisessä on myös vaaransa; julkaisuarkistot ovat pikkuhiljaa täyttymässä automaattisesti haravoidulla metatiedolla, jolla on käyttöarvonsa ylläpitäjien ja julkaisijoiden apuvälineenä tiedon syöttövaiheessa, mutta kokotekstidokumentteja etsivää loppukäyttäjää ne ainoastaan hämmentävät. Tätä silmälläpitäen julkaisemisen työnkulkua tultaneenkin kehittämään radikaalisti siten, että julkisesti selailtaviksi tarjotaan ainoastaan viitteitä joihin on liitetty jokin relevantti tiedosto. Yhä useampi taho maailmalla onkin päätenyt ratkaisuun, jossa käsittelemätön metatieto säilötään omaan arkistoonsa ja vasta tarkastettu ja kokotekstin sisältävä viite julkaistaan kaikkien näkyville [RMS07, WN03].



Kuva 9: Julkaisuarkiston yhteydet muihin tietokantoihin.

Kävijäseuranta








Digitaalisessa muodossa olevat julkaisut mahdollistavat niiden käytön seuraamisen entistä tarkemmin. Onkin huomattu, että artikkelien lataamisen ja niihin kohdistuvien viitteiden välillä näyttäisi olevan selkeä yhteys, joka saattaa tarjota entistä paremman mittarin artikkelien hyödyllisyyden mittaamiseen [KB06]. Julkiset kävijätilastot lisäävät myös tutkijoiden halukkuutta julkaista materiaaliaan arkistossa [Tan07]. Niinpä julkaisuarkiston käyttö- ja kävijämääristä ovat kiinnostuneet ylläpidon lisäksi myös sinne materiaalia syöttäneet tutkijat ja opettajat. Ne voivat helpottaa myös loppukäyttäjää artikkelin painoarvon tunnistamisessa ja ohjata näin tiedon hakijoita relevantin tiedon ääreen.

Tällä hetkellä DSpace tarjoaa vain perustyökaluja kävijäseurantaan. Näitä ovat mm. yleisimpien hakujen tilastointi sekä suosituimpien viitteiden tarkastelu. Tämä ei kuitenkaan ole riittävää mielekkäiden raporttien tuottamiseksi, missä kokotekstin lataaminen halutaan erotella viitetietojen selaamisesta. Tämän vuoksi monet eri tahot ovat rakentaneet sovelluksia, joiden avulla DSpacen lokitiedoista voidaan kerätä yksityiskohtaista ja reaaliaikaista tietoa. Niiden avulla tuotettu julkaisukohtainen kävijäyhteenvedo sisältää tyypillisesti sen katselu- ja latauskerrat maittain eriteltynä (kuva 10). Ylläpitäjän tarpeisiin lisäosat mahdollistavat erilaisten yhteisö- ja kokoelmakohtaisten yhteenvedojen tuottamisen [RMS07, Pap07].

TDS:n kävijäseuranta on toteutettu usealla toisiaan täydentävällä tavalla, joiden antamia tietoja on mahdollista verrata toisiinsa kokonaiskuvan saamiseksi. Tällä hetkellä pääasiallinen kävijäseuranta tapahtuu Google Analytics -palvelun avulla, joka seuraa sivuston käyttömääriä sivuihin liitetyn seurantatägin avulla. Sekundääristä kävijädataa saadaan analysoimalla palvelimen tuottamia lokitiedostoja WebTrends-sovelluksella. Siinä haasteena on kuitenkin relevantin datan erottelu ja tulkinta kaikista lokiin tallennetuista sivupyynnöistä. Kevään 2008 aikana järjestelmässä pyritään ottamaan käyttöön Munhon yliopistossa erityisesti DSpacea varten kehitetty sovellus, joka korvanee ainakin osan nykyisistä seurantamalleista [RMS07].

Meilahden kampusarkisto on noin vuoden toimintansa aikana kerännyt kävijöitä kiittävästi. Julkaisuarkistossa ja kurssikirjastossa oli marraskuuhun 2007 mennessä

käynyt yksittäisiä vierailijoita noin 12 000 kpl. Myös kävijöiden maantieteellinen jakauma on kiinnostava; järjestelmä näyttää keränneen vierailijoita ympäri maailmaa. Seurantatiedoista on havaittu myös hakukoneiden merkittävä rooli käyttäjien tuomisessa TDS:ään.

Item					
Title : Património geomorfológico : conceptualização, avaliação e divulgação : aplicação ao Parque Natural					
Entry Date : 19-07-2007					
Downloads by country			Views by country		
Origin	Downloads	Perc.(%)	Origin	Views	Perc.(%)
 Portugal	65,2	78,05	 Portugal	240	67,80
 Universidade do Minho	9,1	10,89	 Universidade do Minho	49	13,84
 France	6,2	7,48	 Brazil	27	7,63
 N/A	1,0	1,20	 United States	13	3,67
 United States	1,0	1,20	 France	13	3,67
 Brazil	0,8	0,90	 N/A	3	0,85
 Ukraine	0,2	0,30	 Ukraine	2	0,56
	83,6	100,00	 Canada	2	0,56
			 Finland	1	0,28

Kuva 10: Julkaisuarkiston yhden viitteen julkiset kävijätilastot.

5.4 Järjestelmän tietosisältö

Julkaisuarkistossa materiaali on ryhmitelty hierarkkiseen kokoelmarakenteeseen. DSPacen tietomallin mukaisesti rakenteen tasot muodostuvat yhteisöistä, alayhteisöistä sekä näihin kuuluvista kokoelmista. Näistä ainoastaan kokoelmatasot pitävät sisällään varsinaisia viitteitä, yhteisötasojen luodessa niiden välille loogisia suhteita [SBB03].

Kokoelmarakenne

Kampusarkiston kaksoisrooli julkaisuarkistona ja digitaalisena kurssikirjastona osoittautui haasteelliseksi erityisesti kokoelmarakenteen toteutuksen osalta; palvelujen käyttötarpeet ovat erilaiset, minkä vuoksi niiden materiaali haluttiin ryhmitellä eri tavoin. Rakenteellisten erojen ja hallittavuuden vuoksi palvelut lienee parempi pitää jatkossakin erillisinä. Yhteisenä lähtökohtana oli kuitenkin pyrkimys mahdollisimman selkeään ja kevyeen kokoelmahierarkiaan.

Tavallisesti kampusarkistojen kokoelmarakenne noudattelee laitos- ja osastojakoja. Niihin perustuvien massiivisten rakenteiden ongelmaksi muodostuu usein kuitenkin materiaalin vähäinen määrä ja huono käytettävyys [DC07]. Meilahden kampuksen piiriin kuuluu yli 80 eri laitosta tai osastoa, minkä vuoksi tästä päätettiin luopua heti alkuunsa. Sen sijaan materiaalityyppien mukainen jako osoittautui käyttäjän kannalta luonnolliseksi tavaksi ryhmitellä aineisto; näin kokoelmarakenne pysyy hallittavan kokoisena ja käyttäjäystävällisenä ja materiaali profiloituu selkeästi (kuva 11). Aineiston haettavuus osastoittain ja laitostasolla huomioitiin kuitenkin lisäämällä julkaisujen metatietoihin laitoskoodille oma kenttensä, joiden perusteella materiaalia voidaan myös hakea (taulukko 2).

Digitaalisen kurssikirjaston käyttö poikkeaa edellisestä huomattavasti. Aineistoa haetaan tietyn vuosikurssin ja sen sisätämien aihepiirien perusteella. Näin ollen aiemminkin käytössä ollut lukuvuosi- ja aihekohtainen kokoelmajako oli siihen sopivin ratkaisu.

Molemmista palveluista löytyy myös ns. virtuaalisia kokoelmia, jotka ovat itse asiassa järjestelmään tallennettuja hakuja. Niiden kautta materiaalia voidaan näennäisesti ryhmitellä erilaisiin osakokonaisuuksiin, esim. laitoksittain tai tekijän mukaan, niiden varsinaisesta tallennuspaikasta huolimatta.

Metatietomalli

Kaikki järjestelmään tallennettava materiaali kuvaillaan metatiedolla tehokkaiden hakujen mahdollistamiseksi. Materiaalityyppien erilainen luonne asettaa niiden kuvailulle kuitenkin hyvin erilaisia vaatimuksia. Esimerkiksi oppimateriaalia kuvailtaessa halutaan tallentaa tieto vuosikurssista ja oppiaineesta. Artikkelin kohdalla oleellista on sen sijaan julkaisutiedot, kuten lehden nimi ja numero sekä ISSN. Tämän vuoksi jokaista materiaalityyppiä kohden joudutaan määrittelemään juuri sille sopivat kuvailukentät, ns. metatietomalli (taulukot 2 ja 3). Se asettaa kuvailussa käytettävät minimivaatimukset sekä toimii pohjana käyttäjälle tarjottavien syöttölomakkeiden rakentamisessa. Haluttaessa kuvailua voidaan toki rikastaa muillakin



Kuva 11: TDS julkaisuarkiston kokoelmarakenne.

kuvailukentillä.

Tällä hetkellä paikallisesti syötettävä materiaali voidaan jakaa kolmeen ryhmään; tutkimusartikkelit, oppimateriaali ja opinnäytteet. Lukuisten yhteistyötahojen myötä materiaalityyppien määrä tulee lähitulevaisuudessa kuitenkin lisääntymään huomattavasti. Mukaan tulevat tällöin mm. erilaiset sarjajulkaisut, kirjat sekä esitteet ja tilastot.

Kuvailujoukkojen määrittely osoittautui erittäin haastavaksi; se näyttäisikin olevan tasapainoilua syöttölomakkeen täyttämiseen vaaditun vaivan sekä kuvailutiedon myöhempään täydentämiseen käytettyjen resurssien suhteen. Yksinkertaiset kenttärakenteet helpottavat materiaalin kuvailua ja nopeuttavat syöttöprosessia, mutta samalla ne heikentävät tiedon arvoa jatkokäytössä sekä rajoittavat siihen kohdistettavia hakumahdollisuuksia. Lisäksi yhteistoiminnallisuus muiden arkistojen kanssa vaatii standardien noudattamista, minkä vuoksi palvelukohtaisten kenttien ylenmää-

räistä käyttöä on vältettävä. Valmiiden vastaavuustaulukoiden puuttuessa MARC-pohjaisista kirjastojärjestelmistä siirretyn metatiedon muuntaminen Dublin Core -muotoon muodostaa myös oman haasteensa. Informaatikoilta ja tutkijoilta saadun palautteen lisäksi suunnittelussa pyrittiin huomioimaan tuoreet suomalaiset kuvailusuositukset²³. Opinnäytteiden kohdalla tultaneen käyttämään Kansalliskirjaston sitä varten laatimaa suositusta²⁴.

Kuvailutieto	Identifier	Qualifier
Nimeke	title	–
Tekijä	contributor	author
Muut tekijät	contributor	author
Kieli	language	iso
Vastuulaitos	creator	corporateName
Laitoskoodi	identifier	laitoskoodi
URI	identifier	uri
Tiivistelmä	description	abstract
Julkaisutiedot	identifier	citation
ISSN	identifier	issn
Aihe	subject	mesh
Pubmed ID	identifier	pubmedid
Julki ²⁵ luokitus	identifier	JULKI
Tuhti ²⁶	identifier	TUHTI

Taulukko 2. Julkaisuarkistoon tallennetun artikkelin metatietomalli

²³mm. FinnEduMeta, E-thesis, TerveSuomi

²⁴<http://ethesis.helsinki.fi/metadata/>

Kuvailutieto	Identifier	Qualifier
Nimeke	title	–
Tekijä	contributor	author
Kieli	language	iso
Vastuulaitos	creator	corporateName
Laitoskoodi	identifier	laitoskoodi
URI	identifier	uri
Kuvaus	description	–
Tyyppi	type	–
Vuosikurssi	identifier	vuosikurssi

Taulukko 3. Julkaisuarkistoon tallennetun oppimateriaalin metatietomalli

5.5 Peruskäyttö

Julkaisuarkiston merkittävimmät palvelut kaikille käyttäjille ovat tiedon etsimiseen ja selailuun liittyvät toiminnot. Niiden lisäksi palveluun on räätälöity erilaisia lisä-arvopalveluja mm. uutuuksseurantaa ja viitetietojen jatkokäyttöä varten.

Tiedon selaaminen

Arkiston viitteitä voidaan selata monella eri tavalla. Yksinkertaisimmillaan se tapahtuu näpäyttämällä kiinnostavan kokoelman nimeä kokoelmarakenteesta (kuva 11). Tämä vie käyttäjän kyseisen kokoelman etusivulle, jossa on nähtävissä 10 viimeksi julkaistua viitettä nimeketietoineen. Listauksen kaikista kokoelman viitteistä aikajärjestyksessä saa klikkaamalla ”Näytä lisää viitteitä” -linkkiä sivun alalaidassa. Julkaisuaajan lisäksi viitteitä voi selailla myös tekijän, nimekkeen tai aiheen mukaan järjestettynä. Tulokset voidaan rajata kokoelma- tai yhteisötasolle tai tarkastelun kohteeksi voidaan ottaa koko arkiston sisältö.

Viitteitä tai hakutuloksia selailtaessa sen nimeketietojen edessä oleva kuvake ilmaisee siihen liitetyn tiedoston tyyppin (kuva 14). Kuvaketta klikkaamalla tiedosto avataan uuteen selainikkunaan. Nimeketiedot toimivat puolestaan linkkinä, jonka

kautta käyttäjä pääsee tutkimaan julkaisun metatietoja tarkemmin (kuva 12).

Title:	Mutations in amphiphysin 2 (BIN1) disrupt interaction with dynamin 2 and cause autosomal recessive centronuclear myopathy.
Citation:	Nat Genet (5 August 2007)
ISSN:	1061-4036
Author:	Anne-Sophie Nicot, Anne Toussaint, Valérie Tosch, Christine Kretz, Carina Wallgren-Pettersson, Erik Iwarsson, Helen Kingston, Jean-Marie Garnier, Valérie Biancalana, Anders Oldfors, Jean-Louis Mandel, Jocelyn Laporte
Description:	Journal Article
Abstract:	Centronuclear myopathies are characterized by muscle weakness and abnormal centralization of nuclei in muscle fibers not secondary to regeneration. The severe neonatal X-linked form (myotubular myopathy) is due to mutations in the phosphoinositide phosphatase myotubularin (MTM1), whereas mutations in dynamin 2 (DNM2) have been found in some autosomal dominant cases. By direct sequencing of functional candidate genes, we identified homozygous mutations in amphiphysin 2 (BIN1) in three families with autosomal recessive inheritance. Two missense mutations affecting the BAR (Bin1/amphiphysin/RVS167) domain disrupt its membrane tubulation properties in transfected cells, and a partial truncation of the C-terminal SH3 domain abrogates the interaction with DNM2 and its recruitment to the membrane tubules. Our results suggest that mutations in BIN1 cause centronuclear myopathy by interfering with remodeling of T tubules and/or endocytic membranes, and that the functional interaction between BIN1 and DNM2 is necessary for normal muscle function and positioning of nuclei.
Date:	2007-08-13
URI:	http://dx.doi.org/10.1038/ng2086
URI:	http://hdl.handle.net/2455/131541

Kuva 12: Viitteen tiedot julkaisuarkistossa.

Tiedon etsiminen

Tiedon etsimistä varten TDS tarjoaa kolme eri tapaa. Nämä ovat yksinkertainen-, laajennettu- sekä kokotekstihaku. Yksinkertainen haku toimii käyttöliittymän vasemmassa reunassa näkyvän hakukentän kautta ja se kohdistuu viitteiden kaikkiin metatietokenttiin (Liite 2). Laajennettu haku puolestaan mahdollistaa yhden tai useamman valikoidun kuvailukentän mukaisten hakujen tekemisen (kuva 13). Haku- ja tehtäessä voidaan käyttää myös boolean-operaattoreita, kuten AND, OR ja NOT sekä hakea useamman sanan fraaseja. Kaikki järjestelmään tallennetut dokumentit indeksoidaan automaattisesti myös sisältönsä osalta ja sanahaku voidaan ulottaa myös dokumenttien sisältöön ns. kokotekstihaun avulla.

Tarkennettu Haku

Hau kohdistus:
Rajaa hakusi yhteisö tai kokoelmatarolla.

Operaattorit	Hakukenttä	Hakusana
	<input type="text" value="Asiasana"/>	<input type="text" value="cancer"/>
<input type="text" value="AND"/>	<input type="text" value="Asiasana"/>	<input type="text" value="Järvinen, Heikki"/>
<input type="text" value="AND"/>	<input type="text" value="Asiasana"/>	<input type="text"/>

Hakusi "[cancer] AND (Järvinen, Heikki)" tuotti 380 tulosta.

Hakutulokset

Tuloksia sivulla 1-10 of 380 Seuraava sivu

1 2 3 4 ... 38

[Colorectal cancer](#)
Pählman, Lars; Järvinen, Heikki J. (2007-04-05)

[Surveillance in Lynch syndrome](#)
Mecklin, Jukka-Pekka; Järvinen, Heikki J. (2007-04-05)

Kuva 13: Julkaisuarkiston laajennettu haku.

Lisäarvopalvelut

Julkaisuarkiston käyttöliittymään on rakennettu useita toimintoja, joiden tarkoitus on tuoda lisäarvoa käyttäjilleen; esimerkiksi RSS-syötteiden avulla uusien julkaisujen seuraaminen kokoelmittain tai koko arkiston tasolla on helppoa. Lisäksi viitteen kuvailutietoja voidaan helposti viedä muihin sovelluksiin sitä varten rakennettujen pikakuvakkeiden avulla. Uusien toimintojen ohella käyttöliittymään on integroitu myös kirjaston olemassa olevia elektronisia palveluita, joiden näkyvyyttä uuden kampusarkiston toivotaan parantavan entisestään.

Kaikki arkiston lisäarvopalvelut ovat käytettävissä viitteen nimeketietojen perässä näkyvien kuvakkeiden kautta (kuva 14). Kunkin viitteen kohdalla käyttäjälle tarjotaan kuitenkin vain sellaisia toimintoja, joiden käyttämiseksi on olemassa tarvittavat kuvailutiedot. Esimerkiksi PubMed-tietokantahakua ei voi tehdä ellei viitteen metatietoihin ole tallennettu julkaisun PMID-tunnistetta.

Terkossa räätälöityjä toimintoja ovat mm. ostoskori artikkelien maksullista tilausta varten sekä artikkelin viitetietojen tuonti kampuksella käytössä olevaan RefWorks-viitteidenhallintaohjelmaan. Ulkoisia hakupalveluja hyödyntäviä lisäpalveluita ovat PubMed-tietokantahaku, Google Scholar -viittaushaku sekä julkaisutietojen perusteella tehtävä SFX-kirjastohaku (kuva 14). Tulevaisuudessa lisäpalveluiden avulla

voitaneen seurata reaaliajassa Web of Science -tietokannasta haettuja viittausmääriä sekä koota käyttäjäkohtaisia dokumenttikoreja kiinnostaville artikkeleille.



Kuva 14: Viitteen edessä on tiedostotyyppin ilmaiseva kuvake ja sen perässä on linkkejä tarjolla oleviin lisäpalveluihin.

5.6 Tiedon syöttöprosessi

Tiedon syöttöprosessi koostuu käyttäjälle tarjottavien syöttölomakkeiden valinnasta sekä niiden kautta lähetetyn materiaalin käsittelyn osavaiheet määrittelevästä työnkulusta. DSpace tarjoaa tätä varten oman järjestelmänsä, joka rakentuu koelmakohtaisten verkkolomakkeiden ja niihin liitettävien tarkastusaskelten varaan. Niiden muokkaus omiin tarpeisiin on kuitenkin hyvin rajoittunutta. Käytännössä syöttöprosessi on kuitenkin juuri se osa järjestelmää, jota halutaan räätälöidä vastaamaan kulloisiakin tarpeita. Tämän vuoksi monet ovat rakentaneet omia syöttöjärjestelmiään ja ohittaneet Dspacen tarjoaman valmiin kaluston. Myös Meilahden julkaisuarkistoon rakennettiin oma tiedonsyöttöjärjestelmänsä.

Työnkulku

Julkaisuarkistoon syötetyn materiaalin työnkulku määräytyy materiaalityypin perusteella. Artikkelit, oppimateriaali ja opinnäytteet käyvät kukin läpi erilaisen prosessin ennen päätymistään kaikkien näkyville verkkoon (kuva 15).

Kun järjestelmään syötetään uusi artikkeli, se käy läpi yhdistetyn tarkastus- ja hy-

väksymisvaiheen. Materiaalityypistä vastaavat nimetyt vastuuhenkilöt saavat sähköpostitse ilmoituksen uudesta lähetyksestä. Heillä on käytössään selainkäyttöinen hallintanäkymä, josta työjonoon tulleita lähetyksiä voidaan seurata, muokata ja poistaa. Tietojen tarkastuksen ja täydennyksen jälkeen tietue julkaistaan kaikkien näkyville arkistoon (kuva 15, kohta 1).

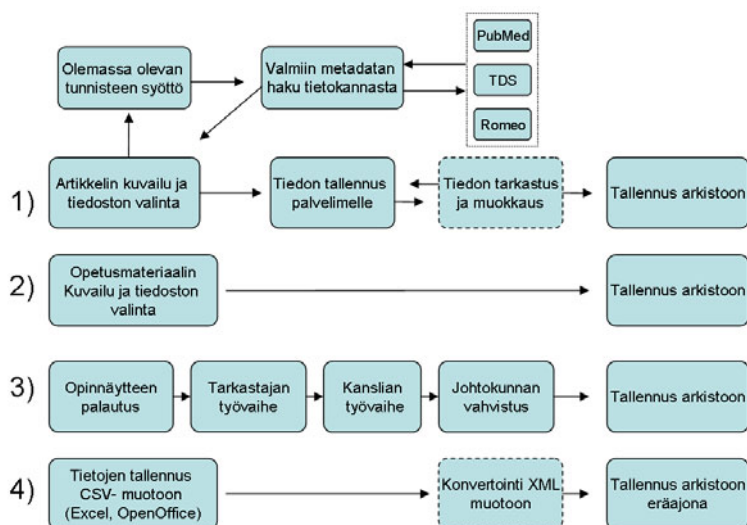
Oppimateriaalin syöttöprosessiin ei sen sijaan liity minkäänlaista tarkastusvaihetta; materiaali näkyy arkistossa välittömästi sen lähetyksen jälkeen. Näin ollen lähettäjällä on täysi vastuu lähettämänsä materiaalin kuvailusta ja laadusta. Järjestely on tiedekunnan opettajien toivomuksen mukainen ja takaa kurssimateriaalin reaaliaikaisen julkaisemisen (kuva 15, kohta 2).

Opinnäytteiden työnkulku on vielä työn alla. Haasteena prosessissa on tiedon kulku eri tahojen välillä sen moniportaisen käsittelyn aikana; asiantuntijatarkastajan ja kanslian jälkeen laitoksen johtoryhmä vahvistaa vielä työn hyväksymisen, mistä ei tavallisesti erikseen enää tiedoteta työn verkkotallennuksesta vastaaville opintoasiainhoitajille. Ideaalitulanteessa julkaisuarkistoa voitaisiin käyttää työkaluna koko tarkastusprosessin eri vaiheiden aikana. Tämä vaatinee kuitenkin pitkäjänteistä kehittystyötä laitoksen opetushenkilökunnan, sekä mahdollisesti yhteistoiminnallisuutta opintorekisterin tietojärjestelmien kanssa (kuva 15, kohta 3).

Syöttölomakkeet

Sisällön saamiseksi järjestelmään tiedon syöttämisen on oltava käyttäjälle helppoa ja nopeaa. Liian pitkät syöttölomakkeet turhauttavat tiedon syöttäjiä ja vähentävät syötetyn materiaalin määrää, liian lyhyet puolestaan heikentävät kerätyn kuvailutiedon laatua ja määrää. Tätä varten jokaista materiaalityyppiä kohden on rakennettu niiden kuvailuun sopivia verkkolomakkeita.

DSpacen oman syöttöjärjestelmän ohittaminen mahdollisti hyvin vapaat kädet lomakkeiden suunnittelussa niin ulkoasun kuin toiminnallisuudenkin suhteen. Lähökohtana oli yksinkertaisuus ja käytön helppous; syöttölomakkeista tehtiin yhden sivun mittaisia, valmiista komponenteista räätälöitäviä kokonaisuuksia, joiden pohjalta uusien lomakkeiden tuottaminen eri yhteisöjen tarpeisiin on helppoa.



Kuva 15: Tiedon syöttöprosessit eri materiaalityypeille: 1) Artikkelin 2) Opetusmateriaali 3) Opinnäyte 4) Eräajona tapahtuva syöttö

Koko syöttöjärjestelmä on toteutettu itsenäisenä PHP-sovelluksena, mikä mahdollistaa järjestelmän käytön myös DSpacesta erillisenä kokonaisuutena. Sen kautta syötetty kuvailutieto tallennetaan XML-muodossa palvelimelle ja voidaan tämän jälkeen syöttää julkaisuarkistoon käyttäen hyväksi sen tarjoamia työkaluja.

Valmiit lomakepohjat on koottu yhdelle sivulle, josta käyttäjä valitsee syöttöprosessin aluksi haluamansa materiaalityypin mukaisen lomakkeen. Valinnan perusteella käyttäjälle tarjotaan materiaalin metatietomallia vastaava lomake täytettäväksi (kuva 16). Lähetettävän materiaalin kuvailutietojen lisäksi lomakkeeseen liitetään myös aineiston tekijän yksikkö- ja laitostiedot, mahdollinen viesti tarkastajalle sekä yksi tai useampia tiedostoja. Lopuksi lähetykseen on mahdollista vielä liittää julkaisuviive (ns. embargo-aika) tai eräpäivä, jolloin viite automaattisesti poistetaan näkyvistä.

Kun lomake on lähetetty, materiaalityypin liitetty työnkulku määrää julkaistaanko sen sisältämät tiedot saman tien vai tallennetaanko ne ylläpitäjän jatkokäsittelyä varten. Onnistuneen prosessin jälkeen lähettäjälle palautetaan vielä tietö viitteen pysyvistä tunnistesta (Handle), jonka avulla siihen on helppo viitata myöhemmin.

Syöttöprosessin helpottamiseksi lomakkeisiin rakennettiin mahdollisuus hyödyntää jo olemassa olevaa metatietoa. Käyttäjä voikin syöttää lomakkeeseen PubMed-tietokannan tai jonkin julkaisuarkiston Handle-tunnisteen, jolloin syöttölomake esitätetään tietokannasta haetulla valmiilla kuvailutiedolla. Tämän jälkeen lomaketta voidaan vielä täydentää tavalliseen tapaan ennen sen lähettämistä.

Lomakkeiden käytössä myös tutkijoita kiinnostava tekijänoikeusnäkökulma on huomioitu. Syötetyn ISSN-numeron perusteella käyttäjälle haetaan Romeo-tietokannasta tiedot kyseisen lehden kustantajan tekijänoikeuskäytännöistä (kuva 16). Näin tiedon syöttäjä voi varmistua oikeudestaan artikkelin verkkotallennukseen jo syöttövaiheessa, eikä materiaalin arkistointi jää tekemättä ainakaan epätietoisuuden vuoksi [SB05].

Kuvailutiedot	
* Title	Cocaine produces fine structural nucle... Acoustic rhinometry.
Abstract	NG 108-15 neuroglioblastoma cells wei
Citation	Brain Res Bull: 20 (1). 39-47, Bull Environ Contam Toxicol, Vol. 76 2006), pp. 497-504.
ISSN	0361-9230
* Language	English
Subject	Animals; Cell Line; Cell Nucleus; Cocai
PubMed ID	3342342
* Institute	Haartman-instituutti (Kliinisteoreettinen laitos)
* Department	Virologian osasto (Kliinisteoreettinen laitos. Haartman-instituutti)
Tiedoston valinta ja tekijänoikeudet	
Lehden tekijänoikeustiedot (Romeo)	
Kustantaja	Elsevier Science B.V., Amsterdam.
Lehti	Brain Research Bulletin
Preprints	restricted
Restrictions	
Postprints	can
Postrestrictions	
Romeo Colour	green
Conditions	On authors personal or authors institutions server Published source must be acknowledged Must link to journal home page Publishers version/PDF cannot be used Articles in some journals can be made Open Access on payment of additional charge NIH Authors articles will be submitted to PMC after 12 months.
Copyright policy	view policy

Kuva 16: Artikkelin syöttölomake tekijänoikeustietoineen.

Suurten tietomäärien syöttö

Terkossa on kehitetty oma sovellus myös suurten tietomäärien syöttämistä varten. Sellainen on koettu erittäin tarpeelliseksi esim. etätyönä tai ilman verkkoyhteyttä työskenneltäessä. Se on osoittautunut myös korvaamattomaksi työkaluksi kampusarkistojen tai uusien kokoelmien perustamisvaiheessa. Tällöin siirron kohteena on kokonaisia aineistotietokantoja, jotka sisältävät tuhansia tietueita.

Sovellukselle annetaan syötteenä jollain taulukkolaskentaohjelmalla tuotettu CSV-tiedosto, jossa jokainen rivi muodostaa yhden kuvailutietueen ja sarake metatietokentän. Kuvailutiedon lisäksi riveille voidaan lisätä viitteeseen liitettävän tiedoston paikallinen polku tai verkko-osoite. Tämän jälkeen PHP-sovellus purkaa tiedoston sisältämät rivit hakemistorakenteeseen, jossa jokaista riviä vastaa yksi alihakemisto. Viitekohtaiset kuvailutiedot kootaan XML-muotoiseen tiedostoon ja rivillä määritellyt tiedostot haetaan automaattisesti samaan hakemistoon. Tämän jälkeen koko hakemistorakenne tiedostoineen voidaan syöttää kerralla sisään DSpacen omien työkalujen avulla.

Viitteiden sarjasyöttö on mahdollista myös PubMed-tietokannasta. Edellisestä poiketen ohjelmalle välitetään kootusti joukko tietokannan PMID-tunnisteita, joita vastaavat metatiedot se hakee tietokannasta ja tallentaa XML-muodossa sisääntuontia varten. Vastaavia sovelluksia viitteiden sisääntuonnin ja automaattisen keräämisen helpottamiseksi on kehitelty myös mm. Singaporen yliopistossa [Tan07].

5.7 Käyttötapaukset

Järjestelmän käyttäjät voidaan jakaa kolmeen ryhmään; tiedon selaajat, materiaalin syöttäjät sekä ylläpitäjät. Ensimmäiseen ryhmään voi kuulua kuka tahansa opiskelijoista yliopiston henkilökuntaan tai satunnaiseen vierailijaan joka selailee julkaisuarkistoa tai kohdistaa sinne hakuja. Materiaalin syöttäjien joukkoon profiloituvat pääasiassa tutkijat ja opettajat sekä asiakasorganisaatioiden vastuuhenkilöt, jotka syöttävät materiaalia järjestelmään. Myös kirjaston henkilökunta voi toimia tässä roolissa syöttäessään materiaalia toisten puolesta. Ylläpitäjiin lukeutuvat järjestel-

män tekniset kehittäjät sekä kirjaston nimeämät muut vastuuhenkilöt, jotka voivat lisätä ja poistaa käyttäjiä sekä hallinnoida oman vastualueensa kokoelmien sisältöjä ja käyttöoikeuksia.

Nykyisellään palvelun käyttötapaukset ovat pääasiassa yksisuuntaisia ja liittyvät syötetyn materiaalin etsintään tai siihen viittaamiseen tunnisteiden avulla. Tulevaisuudessa julkaisuarkistojen ympärille on kuitenkin tarkoitus rakentaa vuorovaikutteisuutta tukeva tutkimus- ja oppimisympäristö, jolloin siihen lisätään välineitä mm. vertaisarviointiin ja ryhmätyöskentelyyn.

Alla on hahmoteltu neljän erityyppisen käyttäjän mahdollisia käyttötapauksia järjestelmässä. Tapaukset on numeroitu relevanssin mukaan siten, että ensimmäisenä luetellaan käyttäjän kannalta olennaisimmat tavat käyttää järjestelmää.

Tutkija

1. Artikkelin tai tutkimusdatan syöttäminen arkistoon
2. Artikkeleiden tai materiaalin etsiminen hakutoimintojen avulla
3. Oman tutkijaportfolion luominen tai täydentäminen arkistossa
4. Uutusseuranta RSS-syötteiden avulla
5. Omien artikkelien kävijämäärien seuraaminen
6. Aineiston saatavuuden selvittäminen SFX-linkin kautta
7. Viitteiden poiminta RefWorks-viitteidenhallintaohjelmaan
8. Painetun aineiston tilaus
9. Tulevaisuudessa: artikkelin reaaliaikaisten viitetietojen seuraaminen arkiston kautta
10. Tulevaisuudessa: osallistuminen tiedonrakenteluun, opintopiireihin tai tutkimusyhteistyöhön järjestelmän tarjoamien työkalujen avulla

Opettaja

1. Uuden materiaalin syöttäminen tai materiaalin päivittäminen arkistoon ja sen tunnisteiden jakaminen opiskelijoille
2. Olemassa olevan materiaalin tai artikkelin hyödyntäminen opetuksessa
3. Oman oppimateriaalin käyttömäärien seuraaminen
4. Uutusseuranta RSS-syötteiden avulla
5. Tulevaisuudessa: verkkokurssien ohjaaminen ja niihin osallistuminen järjestelmän tarjoamien, sosiaalista mediaa tukevien työkalujen avulla

Opiskelija

1. Artikkeleiden tai oppimateriaalin etsiminen hakutoimintojen avulla
2. Materiaalin lukeminen saadun tunnisteiden (Handle) avulla
3. Kurssi tai lukuvuosikohtainen uutusseuranta RSS-syötteiden avulla
4. Artikkelien tai muun aineiston saatavuuden selvittäminen SFX-linkin kautta
5. Artikkelin viitteiden poiminta RefWorks-viitteidenhallintaohjelmaan
6. Tulevaisuudessa: omien harjoitustöiden ja aineistokokoelmien rakentaminen sekä niiden jakaminen muille
7. Tulevaisuudessa: Osallistuminen opintopiirityöskentelyyn ja tiedonrakenteluun sekä omien töiden julkaiseminen järjestelmän tarjoamien työkalujen avulla

Muu käyttäjä

1. Materiaalin selailu ja tiedon etsiminen hakutoimintojen avulla
2. Uutusseuranta RSS-syötteiden avulla
3. Aineiston saatavuuden selvittäminen SFX-linkin kautta
4. Aineiston tilaus
5. Tulevaisuudessa: omaehtoinen-, informaali oppiminen sekä osallistuminen avoimiin opintopiireihin tms.

5.8 Jatkokehitys

Julkaisuarkiston kehitys jatkuu vääjäämättä. Jatkossa tultaneen paneutumaan tarkemmin kampustasolla toimivien julkaisuarkistojen yhteisen hakunäkymän toteuttamiseen sekä yhtenäisten syöttöprosessien ja työnkulkujen kehittämiseen koko HY:n sisällä. Toinen merkittävä haaste on Kansalliskirjaston hallinnoiman URN-tunnusjärjestelmän käyttöönotto ja integroiminen osaksi Dspace-ohjelmistoa ja sen myötä kaupallisista Handle-tunnuksista luopuminen.

Julkaisuarkistojen näkyvyyden esiintuomiseksi ne on saatava kommunikoidaan sujuvasti myös olemassa olevien kirjastojärjestelmien kanssa. Yhteinen SRW/U-rajapinta mahdollistaisi arkistojen sisällyttämisen osaksi kansallista Nelli-portaalia ja sen kautta tehtyjä hakuja. Tämä näyttäisi olevan jo mahdollista koekäytössä olevan lisäosan avulla, joskin vain rajattujen kuvailukenttien osalta.

Kampusarkistoista on kaavailtu seuraajaa yliopiston julkaisurekisteri Julkille. Uudessa mallissa laitosarkistoihin tallennettu tieto toimisi raakadatana Julki-tallentajille, jotka tarkentaisivat ja täydentäisivät sitä tarpeen mukaan. Tällä tavoin julkaisurekisteri ja kampusarkistot sulautuisivat yhteen, mikä puolestaan toimisi mandaattina avoimen verkkotallennuksen puolesta. Tekniset edellytykset järjestelylle ovat jo pitkälti olemassa. Selvityksen alla ovat vielä työnkulkuun liittyvät asiat sekä Dublin Core ja MARC -formaattien väliseen konversioon liittyvät haasteet. Järjestelmien integroiminen edellyttää myös koko HY:n yhteisen metadatatmallin kehittämistä ja käyttöönottoa.

Tulevaisuudessa kampusarkistoissa pyritään ottamaan käyttöön keskitetty käyttäjätunnistus Shibboleth- ja LDAP-tekniikoita hyväksikäyttäen. Järjestely mahdollistaisi DSpace:n omista tunnuksista luopumisen ja yliopiston olemassaolevien tunnus-ten hyödyntämisen käyttöoikeuksien hallinnassa. Suomen virtuaaliyliopisto -hankkeen puitteissa tällaista on jo pilotoitu ja siitä saatujen kokemusten myötä tekniikka voitaneen valjastaa käyttöön myös Helsingin yliopistossa.

Julkaisuarkistojen potentiaali opetuskäytössä on ilmeinen. Käytännön sovellukset sen hyödyntämiseksi ovat kuitenkin vasta ideointivaiheessa. Jatkossa TDS:ään pyritään integroimaan sosiaalista mediaa ja yhteistoiminnallisuutta tukevia toimin-

nallisuuksia, esim. Sakai-ohjelmiston²⁷ avulla.

6 Pohdintaa

Avoim verkkojulkaiseminen on tulossa vahvasti mukaan tieteelliseen julkaisukulttuuriin. Tähän mennessä julkaisuarkistojen tekniikka ja filosofia ovat saaneet julkisessa keskustelussa paljon huomiota. Niiden sijaan arkistoja ylläpitävien yhteisöjen olisi nyt keskityttävä laadukkaan sisällön hankintaan sekä arkistojen profiilin nostamiseen; vasta niiden kautta arkistot saavat ansaitsemaansa näkyvyyttä ja nousevat merkittäviksi toimijoiksi osana korkeakoulujen ja koko yhteiskunnan osaamis pää omaa.

Tällä hetkellä julkaisuarkistojen rakentaminen ja sisällönhankinta tuottavat lisätyötä ja päällekkäisyyksiä niitä ylläpitäville organisaatioille. Tiedon avoimuus ei kuitenkaan välttämättä vaadi erillistä teknistä infrastruktuuria; toimivaan ratkaisuun voidaan päästä myös päivittämällä olemassa olevat järjestelmät vastaamaan ajan haasteita ja mahdollistamaan halutun tiedon julkinen saatavuus. Tällöin avoimuudesta voitaisiin puhua julkaisujärjestelmän ominaisuutena erillisen avoimen arkiston sijaan [Hil07]. Tällä välttyttäisiin useamman rinnakkaisen julkaisujärjestelmän ylläpitämiseltä ja metatiedon siirtoon ja normalisointiin liittyviltä ongelmilta. Mikä parasta, näin myös itsearkistoinnin tuomasta viiveestä ja lisätyöstä voitaisiin päästä eroon.

Avoimien tietovarastojen rakentamisen ohella tulisi huomioida myös niiden mielekkäät käyttömahdollisuudet erilaisissa oppimisympäristöissä ja opetustilanteissa. Oppimisprosessin valjastaminen hyödyntämään niitä lieneekin seuraava suuri haaste opetushenkilökunnalle. Oppilaitosten tulisikin kehittää käytäntöjä oppimateriaalin järjestelmällistä verkkotallennusta silmälläpitäen. Opettajien tulisi puolestaan rohkeammin jakaa tekemäänsä materiaalia näiden uusien väylien kautta sekä opetella ennakkoluulottomasti hyödyntämään julkaisuarkistoja ja niiden tarjoamia mahdollisuuksia osana oppimisprosessia.

²⁷<http://sakaiproject.org/>

Kuluneen 2 vuoden aikana rakennettu julkaisuarkistoympäristö on ollut monella tapaa haasteellinen projekti. Sen myötä avoimen lähdekoodin DSpace on osoittanut suoriutuvansa varsin hyvin avoimen julkaisemisen haasteista. Tarjoamiensa rajapintojen ja työkalujen avulla järjestelmä soveltuu varsin mainiosti myös oppimateriaalin jakeluun. Puuttumaan jää kuitenkin tuki vuorovaikutteisuutta luoville sovelluksille, jotka laajemmassa opetuskäytössä ovat välttämättömiä. Julkaisuarkistosovellusten onkin vielä kehityttävä ennen kuin ne voivat korvata jo käytössä olevat oppimisympäristöt.

Lähteet

- AK07 Altman, M. ja King, G., A proposed standard for the scholarly citation of quantitative data. *D-Lib Magazine*, 13,3/4(2007). URL <http://www.dlib.org/dlib/march07/altman/03altman.html>.
- ANS02 Ahlborn, B., Nejdl, W. ja Siberski, W., OAI-P2P: a peer-to-peer network for open archives. *Parallel Processing Workshops, 2002. Proceedings. International Conference on*, sivut 462–468. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1039766.
- ARL00 Principles for emerging systems of scholarly publishing. *Association of American Universities and Association of Research Libraries and the Merrill Advanced Studies Center of the University of Kansas, 2000*. URL <http://www.arl.org/bm~doc/tempe1.pdf>. Yhteinen julkilausuma.
- Arm00 Arms, W., Digital libraries for distance education. *D-Lib Magazine*, 6,10(2000). URL <http://www.dlib.org/dlib/october00/10editorial.html>.
- Bar06 Barney, C. W. J., What is open access? *Teoksessa: Jacobs, N., Open Access: Key Strategic, Technical and Economic Aspects, Chandos Publishing, 2006*, sivut 13–26.
- BOA02 Budapest open access initiative, 2002. URL <http://www.soros.org/openaccess/read.shtml>. Liikkeen julkilausuma.
- Bro04 Brody, T., Citation analysis in the open access world. *Interactive Media International, 2004*. URL <http://eprints.ecs.soton.ac.uk/10000/>.
- BW05 Barton, M. R. ja Waters, M. M., *Creating an Institutional Repository: LEADIRS Workbook*. MIT Libraries, 2004-2005. URL <http://www.dspace.org/implement/leadirs.pdf>.

- CK04 Capon, N. ja Kuhn, D., Whats so good about problem based learning? *Cognition and Instruction*, 22,1(2004), sivut 61–79.
- Coc06 Cockerill, M., Business models in open access publishing. *Teoksessa: Jacobs, N., Open Access: Key Strategic, Technical and Economic Aspects, Chandos Publishing, 2006*, sivut 111–119.
- Cro02 Crow, R., SPARC institutional repository checklist & resource guide. Tekninen raportti, The Scholarly Publishing & Academic Resources Coalition (SPARC), 2002. URL http://www.arl.org/sparc/bm~doc/IR_Guide_&_Checklist_v1.pdf.
- DC07 Davis, P. M. ja Connolly, M. J. L., Institutional repositories evaluating the reasons for non-use of cornell universitys installation of dspace. *D-Lib Magazine*, 13,3(2007). URL <http://www.dlib.org/dlib/march07/davis/03davis.html>.
- DeR04 DeRidder, J., Choosing software for an institutional repository. Tekninen raportti, UTK Libraries Digital Library Center, 2004. URL <http://aztec.lib.utk.edu/~deridder/comparison.doc>.
- DHM04 Dziuban, C. D., Hartman, J. L. ja Moskal, P. D., Blended learning. *Educause Center for Applied Research Bulletin, Issue 7, 2004*. URL <http://ecampus.nl.edu/resources/webct/BlendedLearningUCFStudy.pdf>.
- EC06 European commission: Study on the economic and technical evolution of the scientific publication markets in Europe, January 2006. URL http://ec.europa.eu/research/science-society/pdf/scientific-publication-study_en.pdf.
- Ers06 Erstad, O., A new direction? Digital literacy, student participation and curriculum reform in Norway. *Education and Information Technologies*, 11, sivut 415–429.

- GA08 Green, R. ja Awre, C., RepoMMan: Delivering private repository space for day-to-day use. *Ariadne*, 54, 2008. URL <http://www.ariadne.ac.uk/issue54/green-awre/>.
- Hak02 Hakala, J., Metadata: käyttökohteet ja formaatit. *Teoksessa: Haasio, A. ja Piukkula, J., Tietoverkot ja kirjastot, BTJ Kirjastopalvelu, 2002.*
- Har06 Harnad, S., Opening access by overcoming zenos paralysis. *Teoksessa: Jacobs, N., Open Access: Key Strategic, Technical and Economic Aspects, Chandos Publishing, 2006*, sivut 73–85.
- HBV⁺04 Harnad, S., Brody, T., Vallieres, F., Carr, L., Hitchcock, S., Gingras, Y., Oppenheim, C., Stamerjohanns, H. ja Hilf, E. R., The access/impact problem and the green and gold roads to open access. *Serials Review*, 30,4(2004), sivut 310–314. URL <http://www.sciencedirect.com/science/article/B6W63-4DS8FN9-F/2/8604ea015ea3f8c71f055271f050255a>.
- Hil07 Hilska, K., Julkaisutietojärjestelmä julia KTL:n avoimena arkistona. *Julkaisuarkistot Workshop Helsingin Yliopistolla 11.1.*, 2007.
- HLI99 Hakkarainen, K., Lipponen, L., Ilomäki, L. et al., *Tieto ja viestintätekniikka tutkivan oppimisen välineenä*. Helsingin kaupungin opetusvirasto, Tietotekniikkaprojektin tutkimusryhmä, Helsinki, 1999. URL http://www.helsinki.fi/science/networkedlearning/julkaisut/tvt_tutkivan_oppimisen_valineena.pdf.
- HLL99 Hakkarainen, K., Lonka, K. ja Lipponen, L., *Tutkiva oppiminen: Älykkään toiminnan rajat ja niiden ylittäminen*. WSOY, 1999.
- HSG05 Hajjem, C., Stevan, H. ja Gingras, Y., Ten-year cross-disciplinary comparison of the growth of open access and how it increases research citation impact. *IEEE Data Engineering Bulletin*, 28,4(2005), sivut 39–47. URL <http://eprints.ecs.soton.ac.uk/11688/>.

- Jak07 Jakobsson, A., NORA (norwegian open research archives) and the development of institutional repositories in Norway. *Julkaisuarkistot Workshop Helsingin Yliopistolla 11.1.*, 2007. URL <http://www.openaccess.fi/info/julkark07/jakobsson.ppt>.
- JAM06 Jones, R., Andrew, T. ja MacColl, J., *The Institutional Repository*. Chandos information professional series. Chandos Publishing, Oxford, England, 2006.
- KB06 Kurtz, M. ja Brody, T., The impact loss to authors and research. *Teoksessa: Jacobs, N., Open Access: Key Strategic, Technical and Economic Aspects, Chandos Publishing, 2006*, sivut 45–54.
- KEA⁺05 Kurtz, M. J., Eichhorn, G., Accomazzi, A., Grant, C., Demleitner, M., Henneken, E. ja Murray, S. S., The effect of use and access on citations. *Information Processing & Management*, 41,6(2005), sivut 1395–1402. URL <http://www.sciencedirect.com/science/article/B6VC8-4G05MJJ-1/2/db2076becb4d9121ba6c6330b9e4e439>.
- Kes07 Keskitalo, E.-P., Meillä on vain nimet : URN-tunnukset helpottavat tiedonhakua. *Tietolinja*, 1,1(2007). URL <http://urn.fi/URN:NBN:fi-fe20071298>.
- Ket07 Kettunen, R., Open access - tieteellisen tiedon avoin saatavuus. *Tieteessä Tapahtuu*, 25,7(2007). URL <http://ojs.tsv.fi/index.php/tt/article/viewFile/320/283>.
- Kru06 Krumsvik, R., The digital challenges of school and teacher education in Norway: Some urgent questions and the search for answers. *Education and Information Technologies*, 11,3-4(2006), sivut 239–256.
- Law01 Lawrence, S., Online or invisible? *Nature* 411(6837):521, 2001.
- LJR05 Levonen, J., Joutsenvirta, T. ja R., P., Blended learning - katsaus sulautuvaan yliopisto-opetukseen. *Piirtoheitin*, 3,2(2005). URL <http://www.valt.helsinki.fi/piirtoheitin/sulautus1.htm>.

- Lom06 Lomangino, K., *Institutional Repositories: Their Emergence and Impact on Scholarly Publishing*. The Sheridan Press, 2006. URL http://www.sheridanpress.com/assets/pdf/inst_repositories.pdf.
- LVdS07 Lagoze, C. ja Van de Sompel, H., March 3 release of OAI-ORE specifications. Tekninen raportti, Open Archives Initiative, 2007. URL <http://www.openarchives.org/ore/documents/ore-hopkins-press-release.pdf>.
- Lyn01 Lynch, C. A., Metadata harvesting and the open archives initiative. *ARL: A Bimonthly Report*, no. 217, 2001, sivut 1–9. URL <http://www.arl.org/resources/pubs/br/br217/br217mhp.shtml>.
- MB07 Mornati, S. ja Bollini, A., A new opportunities for institutional repositories: the evaluation challenge, a case study. *Presentation at DSpace User Group Meeting, Rome, Italy*, 2007. URL <http://www.aepic.it/conf/viewabstract.php?id=214&cf=11>.
- McC99 McCabe, M. J., The impact of publisher mergers on journal prices: An update. *ARL: A Bimonthly Report*, no. 207, 1999, sivut 1–5. URL <http://www.arl.org/resources/pubs/br/br207/br207jrnlprices.shtml>.
- MCNB05 McCown, F., Chan, S., Nelson, M. L. ja Bollen, J., The availability and persistence of web references in d-lib magazine. *5th International Web Archiving Workshop, Vienna, Austria*, 2005, URL <http://arxiv.org/abs/cs/0511077v1>.
- MM95 Marchionini, G. ja Maurer, H., The roles of digital libraries in teaching and learning. *Commun. ACM*, 38,4(1995), sivut 67–75.
- Nii01 Niiniluoto, I., Julkisuusperiaate ja tutkimustulosten salailu. *Tieteesä Tapahtuu*, 19,2(2001). URL <http://www.tsv.fi/ttapaht/012/niiniluoto.htm>.

- Ock07 Ockerbloom, J. M., Dspace 2.0 architecture review report. *Open Repositories 2007 conference, San Antonio, TX*. DSpace architecture review group, January 2007, URL http://wiki.dspace.org/static_files/0/0e/DSpace-recs.pdf.
- Ope05 Opetusministeriö, Avoimen tieteellisen julkaisutoiminnan työryhmän muistio, 2005. URL http://www.minedu.fi/OPM/Julkaisut/2005/avoimen_tieteellisen_julkaisutoiminnan_tyoryhman_muistio.
- Pap07 Paparoni, J., Statistics for dspace. *Presentation at DSpace User Group Meeting, Rome, Italy*, 2007. URL <http://www.aepic.it/conf/viewpaper.php?id=337&cf=11>.
- PGL05 Phillips, S., Green, C. ja Legget, J. e. a., Manakin: Developers guide. Tekninen raportti, Digital Initiatives research & technology, Texas A&M Library, 2005. URL <http://di.tamu.edu/projects/xmlui/resources/DevelopersGuide.pdf>.
- PGM07 Phillips, S., Green, C. ja Maslov, Alexey, e. a., Manakin - a new face for dspace. *D-Lib Magazine*, 13,11(2007). URL <http://www.dlib.org/dlib/november07/phillips/11phillips.html>.
- Pri04 Pringle, J., Do open access journals have impact?, *Nature (Web Focus)*, 2004. URL <http://www.nature.com/nature/focus/accessdebate/19.html>. Kolumni.
- Pro07 Prosser, D., Open access repositories: Policy and practice, SPARC, Europe. *Presentation at Jag trivs bäst i öppna landskap conference*, 2007. URL <http://www.ub.uio.no/nora/landskap/presentation/sparc.ppt>.
- RMS07 Rodrigues, E., Miranda, A. ja Saraiva, R., Improving usage statistics for repositório. *Presentation at DSpace User Group Meeting, Rome, Italy*, 2007. URL <http://www.aepic.it/conf/viewpaper.php?id=209&cf=11>.

- Roe01 Roes, H., Digital libraries and education trends and opportunities. *D-Lib Magazine*, 7,7/8(2001). URL <http://www.dlib.org/dlib/july01/roes/07roes.html>.
- Saa07 Saariluoma, P., Open access publishing as a bridge across the digital divide. *An Interdisciplinary journal on Human in ICT Environments*, 3 (2007), sivut 116–119.
- Sal06 Sale, A., Researchers and institutional repositories. *Teoksessa: Jacobs, N., Open Access: Key Strategic, Technical and Economic Aspects, Chandos Publishing, 2006*, sivut 87–98.
- SB05 Swan, A. ja Brown, S., Open access self-archiving: An author study, Key Perspectives Limited, May 2005. URL <http://cogprints.org/4385/>. Kyselytutkimus.
- SBB03 Smith, M., Barton, M. ja Bass, Mick, e. a., Dspace-an open source dynamic digital repository. *D-Lib Magazine*, 9,1(2003). URL http://dspace.ibict.br/dmdocuments/Dspace_%20An_Open_Source_Dynamic_Digital_Repository.pdf.
- SBCH06 Shadbolt, N., Brody, T., Carr, L. ja Harnad, S., The open research web. *Teoksessa: Jacobs, N., Open Access: Key Strategic, Technical and Economic Aspects, Chandos Publishing, 2006*, sivut 195–208.
- Sch03 Schroeder, K., The application of persistent identifiers as one approach to ensure guarantee long-term availability of online-theses; the established uniform resource name (URN) -management at die deutsche bibliothek. *Presentation at Electronic Theses and Dissertations Worldwide conference*, Berlin 2003.
- SGWG06 Sorokina, D., Gehrke, J., Warner, S. ja Ginsparg, P., Plagiarism detection in arXiv. *ICDM '06. Sixth International Conference on Data Mining*, sivut 1070–1075, 2006

- Sub07 Suber, P., Open access overview: Focusing on open access to peer-reviewed research articles and their preprints. *The SPARC Open Access Newsletter*, June 2007. URL <http://www.earlham.edu/~peters/fos/overview.htm>.
- SW04 Sompel, L. ja Warner, N., OAI-PMH version 2.0 specification. Tekninen raportti, Open Archives Initiative, 2004. URL <http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm>.
- Swa06 Swan, A., The culture of open access: researchers views and responses. *Teoksessa: Jacobs, N., Open Access: Key Strategic, Technical and Economic Aspects*, Chandos Publishing, 2006, sivut 66–72.
- Swa07 Swan, A., The business of digital repositories, In Press, 2007. URL <http://eprints.ecs.soton.ac.uk/14455/>.
- Tan07 Tansley, R., Dspace architecture: an update. *Presentation at DSpace User Group Meeting, Rome, Italy*, 2007. URL <http://www.aepic.it/conf/viewpaper.php?id=328&cf=11>.
- TM98 Twigg, C. ja Miloff, M., The global learning infrastructure. *Teoksessa: Tapscott, Lowry ja Ticoll: Blueprint to the digital economy*, 1998.
- VdSNL04 Van de Sompel, H., Nelson, M. L., Lagoze, C. et al., Resource harvesting within the OAI-PMH framework. *D-Lib Magazine*, 10,12(2004). URL <http://www.dlib.org/dlib/december04/vandesompel/12vandesompel.html>.
- Wil06a Willinsky, J., *The access principle: the case for open access to research and scholarship*. Cambridge, MIT Press, 2006. URL <https://www.tcrecord.org/Content.asp?ContentID=12338>.
- Wil06b Willinsky, J., Access to power: Research in international policymaking. *Harvard International Review*, 28,2(2006), sivut 54–57. URL <http://pkp.sfu.ca/files/HIR.pdf>.

- WN03 Warner, S. ja Nelson, M., Report on the metadata harvesting workshop at Joint Conference on Digital Libraries (JCDL). Tekninen raportti, 2003. URL http://www.cs.cornell.edu/people/simeon/workshops/JCDL2003/JCDL2003_OAI_Workshop.pdf.
- Wyl06 Wyles, R., Technical evaluation of selected open source repository solutions in open access repositories in New Zealand. Tekninen raportti, CPIT, 2006. URL <https://eduforge.org/docman/view.php/131/1062/Repository%20Evaluation%20Document.pdf>.
- YK03 Geneva declaration of principles, world summit on the information society, December 2003. URL <http://www.itu.int/wsis/docs/geneva/official/dop.html>.

Liite 1. CDSWaren ja DSpacen vertailu

2.3.2006

Vaatimuslista: Kimmo K, Teo K, Kalle K, Jani S
taulukoinut Kalle K, täydentänyt Mika S

CDSWare 0.7 DSpace 1.3.2

Käyttöliittymän tulee olla web-pohjainen. [***]	X	X
Kieliversiot ja niiden helppo ylläpitäminen. Suomi, ruotsi, englanti. [**]	X	X
Palvelun käytön tulee olla helppoa. Käyttäjille ei välttämättä tarvitse antaa erityistä käyttäjäkoulutusta. [***]	X	X
Ohjeistukset. Käyttöliittymässä tulee tarjota helposti ja selkeästi ohjeet palvelun käyttöön [***]	X	X
kolmella kielellä. [**]	X	X
Selaaminen yliopiston organisaation mukaan (tutkimusryhmä pikemmin kuin laitos, jos relevantti) [***]	X	X
– aihealueen mukaan (esim. kemia, kardiologia, kieliteknologia) [**]	X	?
– aineistotyyteittäin (esim. opinnäyte, tieteellinen artikkeli, julkaisusarja) [***].	X	?
– julkaisuvuoden perusteella [**]	X	X
Tiedonhaku (search): Palvelu tarjoaa yksinkertaisen hakulaatikon tiedonhakuun hakusanalla. [***]	X	X
Lisäksi edistyneempi haku, jossa hakua voi kohdistaa eri kenttiin [***]	X	X
Tiedonhaun tulee voida kohdistua myös julkaisuarkistossa oleviin kokoteksteihin. [***]	X	X
Hakutulokset tulee esittää selkeästi. Haun jälkeen tulosta tulee voida järjestellä: ilmestymisajankohdan mukaan [***]	X	X
– aakkosittain [***]	X	X
– relevanssi [*]	X	?

Kokoelmanäkymät tutkimusryhmän tai tutkijan julkaisuista [**]	X	X
Loppukäyttäjä näkee kustakin dokumentista kuvailutietoja, joiden avulla hän voi päättää lataako dokumentin vai ei. [***]	X	X
Julkaisuarkistoon tallennettu dokumentti tulee tarjota loppukäyttäjille eri tiedostoformaateissa (HTML, PDF). Dokumentin voi lukea selainohjelmalla. [*]	X	?
Selkeän linkin luominen julkaisuarkistossa olevaan dokumenttiin. Arkistossa olevilla dokumenteilla tulee olla yksiselitteinen, linkattavissa oleva URI-tunnus tai muu identifiointitunniste. [***]	?	X
Palvelun käyttöliittymä tulee olla HY:n muokattavissa. Palvelun ulkoasussa tulee voida käyttää yliopiston yleistä ulkoasua. [**]	X (rajoituksin)	?
Palvelussa voidaan helposti esittää julkaisuarkiston suosituimmat julkaisut [**]	?	?
Palvelussa voidaan helposti esittää julkaisuarkistoon tulleet uusimmat artikkelit [***]	X	X
Palvelussa voidaan helposti esittää ylläpitäjän valikoimat ns. ”Feature” –aineistot [*]	?	?
Palvelun käyttäjille tarjotaan uutisvirtana lista uusimmista dokumenteista RSS-feedinä (RSS 2.0, Atom). Palvelussa tulee voida toteuttaa useampia RSS-syötteitä. [**]	?	?
Lisäksi palvelun käyttäjille voidaan tarjota uutuustiedot sähköpostilla (Alertit) [*]	X	?
Palveluun tulee voida integroida palautejärjestelmä. Ideaalitulanteessa palautejärjestelmällä voidaan saatu palaute ohjata oikealle henkilölle, joka vastaa tietystä palvelun kokonaisuudesta. [***]	?	?
Julkaisuarkistosovelluksen tulee mahdollistaa yksittäisen dokumentin liittäminen julkaisuarkistoon webin kautta. [***]	X	X
Julkaisuarkistosovelluksen tulee mahdollistaa usean dokumentin (tai dokumenttiin liittyvien osien, liitteiden, kuvituksen) liittäminen julkaisuarkistoon webin kautta yhdellä kerralla. [*]	X	?
Julkaisuarkistoon syötetyt dokumentit tulee sijoittaa	X	X

”odotustilaan” tarkistusta varten ennen kuin ne julkaistaan yleisölle. [***]		
Julkaisun tallentamisen tulee olla käyttäjälle selkeä ja ohjattu prosessi, joka toteutuu web-lomakkeella. [***]	X	X
Palvelun kokoelmien perusylläpito on oltava web-pohjainen [***]	X	X
Auktorisoidut päivittäjät (manager) ylläpitävät dokumenttikokoelmia ja vievät aineistoa julkaisuarkistoon. Päivittäjät hyväksyvät ”odotustilassa” olevat julkaisut järjestelmän julkaistavaksi. [***]	X	X
Web-palvelun/sivuston päivityksen tulee olla joustavaa. [***]	X	X
Järjestelmään tulee pystyä viemään kuvailutdataa yliopiston piirissä toimivista kampusarkistoista. Kampusarkistojen käyttämät metadataformaattit voivat vaihdella.	X	–
Julkaisuarkiston tulee voida toimia sekä service providerina että data providerina.	X	–
Kuvailudatan siirtoprotokollana tulee voida käyttää OAI-PMH-standardia (OAI-PMH 2.0). [***]	X	X
Kuvailudatan tallennusformaatti (tehdasasetus)	MarcXML	DC Qual.
Tukee useita kuvailudatan tallennusformaatteja	X	–
Järjestelmä tarjoaa web-pohjaiset välineet kuvailudatan laadunvarmistukseen.	X	X
Järjestelmän avulla tulee voida hallinnoida myös kokotekstejä. Miten kokotekstit syötetään, mihin tallennetaan, miten hallinnoidaan?	X	X
Dokumenttien tiedostoformaattit? Jos tavoitteena että loppukäyttäjän tulee saada teksti käyttönsä eri muodoissa, niin se vaikuttaa tähän. HTML, PDF...[*]	X	?
Järjestelmällä tulee voida tuottaa kokoelmatason pääsivuja/kokoomasivuja. Esimerkiksi sarjajulkaisuilla tulee olla oma ”etusivu” julkaisuarkistossa.	X	X
Järjestelmään tulee voida luoda kokoelmatason kuvaus esim. sarjajulkaisulle, joka yhdistää sarjajulkaisuun liittyvät dokumentit.	X	X
Järjestelmän tulee tarjota tarkkoja käyttötilastoja palvelun	?	?

käytöstä. Tehdyt haut, ladatut artikkelit, web-sivujen yleistilastot... [**]		
Järjestelmä pitää kirjaa palveluun liitetystä dokumenteista. Kokonaismäärä, tilastot kokoelmittain (tiedekunnat, laitokset, kampukset, artikkelit, opinnäytteet, sarjajulkaisut). (Julkaisutilastot) [**]	X	?
Tietokantasovelluksen vaatimukset	MySQL	PostgreSQL (Oracle)
Vasteajat		
OAI-PMH (muut avoimet julkaisuarkistot, tms.) [***]	X	X
SRU/SRW (Nelli, muut portaaliset) [**]	?	?
RSS-feedit	X	X?
Yhteys sisällönhallintaan (RESTfull API)	?	?
Yhteys JULKI-julkaisutietokantaan (Voyager)	X	?
HELKA ja muut kirjastoluettelot	X	?
MUTI, ASTI? [*]	X?	X?
Muut? Aineistojen siirto kansalliskirjaston kokoelmiin.	X (OAI)	X (OAI)
— POCKLEYN TAULUKOSTA LISÄTTY —		
Versioiden hallinta	X	—
Mahdollisuus asettaa pakolliset metadataelementit	?	—
Tiedoston lisääminen metadatatietueeseen	?	—
Tietueen monistaminen	?	—
Mahdollisuus käyttää lyhennettyjä hakutermejä	X	—

Mahdollisuuksien mukaan dokumenteista tulisi olla saatavilla myös tiivistelmä.[**]	X	X
Kokotekstien esikäsittely, mahdollinen konversio, Tiedostojen nimeämiskäytännöt, kokotekstien tallennus siten että yhteys kuvailudataan säilyy (yksiselitteiset tunnisteet)	konversio mahdollinen lisäpalikalla	konversio mahdollinen lisäpalikalla
Tekijänoikeuksien/julkaisuluvan hallinnasta yksinkertaiset ohjeet [***]	X	X

Liite 2. Julkaisuarkiston DSpace-käyttöliittymä

[About DSpace Software](#)

Search DSpace

Advanced Search

[Home](#)

Browse

[Communities & Collections](#)
[Titles](#)
[Authors](#)
[Subjects](#)
[By Date](#)

Sign on to:

[Receive email updates](#)
[My DSpace authorized users](#)
[Edit Profile](#)
[Help](#)
[About DSpace](#)

Terkko Document Space >

DSpace is Live

Welcome to our digital repository of My University research!

More exciting news to appear here.

Search

Enter some text in the box below to search DSpace.

Communities in DSpace

Choose a community to browse its collections.

[Articles \[Metadata and Full Text\]](#) [36513]
[Conferences, Seminars and Posters](#) [26]
[Kansaneläkelaitos \(Kela\)](#) [3]
[Kansanterveyslaitos \(KTL\)](#) [3]
[Kotimaisia Open Access -lehtiä \[Finnish OA Journals\]](#) [9]
[Opinnäytteet \[Theses and Dissertations\]](#) [177]
[Principal Investigator Profiles \[University of Helsinki\]](#) [241]
[Suomalaista lääketieteen historiaa](#) [809]
[Virtuaaliset laitospokoelmat \[Departments and Institutes\]](#) [0]

This is a default installation of DSpace!

It can be extensively configured by installing modified JSPs, and altering the site configuration.

RSS Feeds

[RSS 1.0](#)[RSS 2.0](#)



DSpace Software Copyright © 2002-2007 [MIT](#) and [Hewlett-Packard](#) - [Feedback](#)

Liite 3. Julkaisuarkiston Manakin-käyttöliittymä


Yliopiston etusivulle

Suomeksi

På svenska

In English

Login



HELSINGIN YLIOPISTO

Terveystieteiden keskuskirjasto

Lääketieteellinen tiedekunta

TDS Etusivu

TDS haku

Hae

Tarkennettu haku

Selaa aineistoa

Koko arkisto

Kokoelmat

Nimekkeet

Tekijät

Aiheet

Julkaisuajat

Oma sivu

Kirjaudu

Rekisteröidy

Linkkejä

Terkko Health

Harvester

Terkko Online

FeedNavigator

JournalNavigator

BookNavigator

Lääketieteellinen tiedekunta

Julki

TerkkoOA

FinnOA

OA ohjeita tutkijoille

Sherpa- Authors

Sherpa/RoMEO

DViikki/ KATJA

RSS 1.0

RSS 2.0

Terkko Document Space (TDS)

TDS on Helsingin yliopiston lääketieteellisen tiedekunnan Meilahden kampuksen avoin julkaisuarkisto. Palveluun tallennetaan kampuksella tuotettua kokotekstiaineistoa sekä julkaistujen artikkelien ja dokumenttien metadataa (viitetietoja).

Palvelua ylläpitää [Terveystieteiden keskuskirjasto](#).

Jos haluat tallentaa materiaalia tai julkaisuja arkistoon, ota yhteyttä [ylläpitoon](#) tai voit käyttää suoraan [syöttölomakkeita](#).

Terkko Document Space is the Institutional Repository of the Medical Faculty, University of Helsinki.

Hae arkistosta

Kirjoita haettava teksti.

Hae

TDS:n kokoelmat

Valitse kokoelma.

- Articles [Metadata and Full Text]
- Conferences, Seminars and Posters
- Kansaneläkelaitos (Kela)
- Kansanterveyslaitos (KTL)
- Kotimaisia Open Access -lehtiä [Finnish OA Journals]
- Opinnäytteet [Theses and Dissertations]
- Principal Investigator Profiles [University of Helsinki]
- Suomalaista lääketieteen historiaa
- Virtuaaliset laitoskokoelmat [Departments and Institutes]

Contant Us | Send Feedback